

University of Memphis

University of Memphis Digital Commons

Electronic Theses and Dissertations

2020

Shared Agency and Structural Oppression: The Ethics of Social Reality Management

Christina Friedlaender

Follow this and additional works at: <https://digitalcommons.memphis.edu/etd>

Recommended Citation

Friedlaender, Christina, "Shared Agency and Structural Oppression: The Ethics of Social Reality Management" (2020). *Electronic Theses and Dissertations*. 2547.
<https://digitalcommons.memphis.edu/etd/2547>

This Dissertation is brought to you for free and open access by University of Memphis Digital Commons. It has been accepted for inclusion in Electronic Theses and Dissertations by an authorized administrator of University of Memphis Digital Commons. For more information, please contact khggerty@memphis.edu.

SHARED AGENCY AND STRUCTURAL OPPRESSION:
THE ETHICS OF SOCIAL REALITY MANAGEMENT

by

Christina Friedlaender

A Dissertation

Submitted in Partial Fulfillment of the

Requirements for the Degree of

Doctor of Philosophy

Major: Philosophy

The University of Memphis

August 2020

DEDICATION

To the memory of my maternal grandmother, Olga Frieda Clara Panholzer (née Sonder):
For your library, for deep intergenerational loss, for your posthumous J.D. and unfinished PhD.

Acknowledgements

While academia can often make intellectual work feel like an individual undertaking, this is, of course, a false picture. In this section, I would like to thank and recognize those people (and places!) who have contributed to my intellectual, personal, and professional development.

Portland State University: I am deeply indebted to the Women, Gender, and Sexuality Studies program at Portland State University. Both their commitment to critical pedagogy, experimental education, and their approach to social action as a form of education has significantly shaped who I am both as a scholar and as an educator. Special thank you to Dr. Jamie Ross for her feminist mentorship and for introducing me to feminist philosophy of science and feminist epistemology. I am indebted to Dr. Ross for my intellectual development as a feminist philosopher.

However, as a feminist thinker, I am indebted specifically to Roslyn Farrington (Med) of the Portland State WGSS Department, who passed away in 2014. Roslyn taught the first class I took in women's studies, and showed the power of critical, feminist pedagogy by practicing it with her students. On the first day of class, Roslyn displayed a visualization of Patricia Hill Collin's notion of a 'matrices of oppression,' framing not only the terms and the scope of our class discussions, but effectively shaping my perception of feminism, theory, and social action in long-term ways. Roslyn is the reason I ever stepped foot into a feminist, community organizing meeting in the first place, and I know this is true for many other, now long-term organizers who took her classes.

Thank you also to Dr. Avram Hiller of the Philosophy Department at Portland State University for his support.

Binghamton University: Thank you to the Philosophy Department at Binghamton University – State University of New York. As a social and political philosopher, I am grateful for the rigorous philosophical training I received there. Special thank you to both Dr. Bat-Ami Bar-On and Dr. Lisa Tessman for her mentorship during my time in the program. Thank you also to Dr. Robert Guay, Dr. Max Pensky, and Dr. Anja Karnein.

The University of Memphis: I want to start by expressing deep gratitude to the Philosophy Department at the University of Memphis. As someone who has never taken particular pride in the educational institutions that have housed me, I can say that I am proud to have a Philosophy degree from the University of Memphis. Choosing a graduate program is not an easy decision, and I, without a doubt, chose the best department for the kind of philosopher I am and the work that I do. I am a better philosopher for having been trained in a genuinely pluralist intellectual space and in a department that did not impose one particular model for how to do philosophy.

Thank you to my committee: Dr. Deborah Tollefsen, Dr. Shaun Gallagher, Dr. Lindsey Stewart, and Dr. Luvell Anderson (Syracuse). I am lucky to have had such an incredible group of people overseeing my work. Thank you also to my teaching mentor, David Miguel Gray. I would also like to thank my advisor, Dr. Deborah Tollefsen, in particular. I cannot count the number of times I have bragged to graduate students from other institutions about having won the advisor lottery. Thank you to her for the many hours spent reading drafts, for supporting my weird research while making sure I did not fall into a rabbit hole or go completely awry, and for fielding far too many emails from me over the years. I am at a loss for words for how significantly Dr. Tollefsen has contributed to my professional and philosophical development.

during my time in Memphis. Second, thank you to Dr. Bill Lawson for encouraging my research on the topic of microaggressions and for changing my approach to writing analytic philosophy. Third, thank you to Shaun Gallagher for his continuous support and professional development opportunities. Lastly, thank you to Lindsey Stewart for further shifting my sense of what is possible in professional philosophy. Thank you to all those who have been involved with the University of Memphis chapter of Minorities and Philosophy. Our group's intellectual rigor and solidarity has been a source of joy throughout my time here. Thank you also to the larger graduate student community in the philosophy department, who have been a constant wellspring of support, philosophical debate, and much needed fun.

Additional Acknowledgements: Thank you to Dr. Dan Zahavi and to The Center for Subjectivity Research at The University of Copenhagen in Denmark for generously hosting me as a visiting PhD student. My time there contributed valuably towards my growth as a philosopher and the completion of my dissertation. Thank you to Kendy Hess, Violetta Igneski, and Tracy Isaacs for including my article, "Resisting Oppression Together: Shared Intentions and Unequal Agents," in their anthology, *Collectivity: Ontology, Ethics and Social Justice*, which eventually served as the draft for my prospectus. Thank you also to audiences at The American Philosophical Association and the International Society for Social Ontology for their feedback on early versions of the work contained herein. Thank you to Jessica Pedrosa for her help with editing.

Friends and Family: Finally, I would like to thank those family and friends who have supported me throughout this entire process. Thank you to the following fellow philosophers for their

friendship, support, and for being razor-sharp interlocutors: Tamsin Kimoto, Courtney Miller, John Torrey, Leonie Smith, and Jasper St. Bernard. Thank you my closest friends Julia Spahle and Jessica Pedrosa for keeping me in check and reminding me that there's a whole world outside of academia. Special thank you to fellow philosopher Joshua Felix for being my closest interlocuter for the last six years and for his continuous and profound support throughout this process. Lastly, I would like to thank: my maternal grandfather, Rudolph Panholzer, and his partner, Danielle Kuska, my parents, Veronika Hertha Panholzer-Doyle and James Jay Warne, and my little brother Michael 'Mikal' Sonder Doyle for their unconditional support and for making all of this possible for me.

ABSTRACT

My dissertation offers a methodological, ideological, and ethical critique of literature on shared agency through the lens of oppression theory. Oppression theory starts from the claim that structural forms of oppression play a significant role in shaping our social world, including our identities, epistemic standpoints, values, and metaphysical worldviews. Connecting Hans Vaihinger's work on methodology to the tradition of feminist philosophy of science, I argue that we develop models of shared agency not merely for the purpose of understanding but also to better manage how we undertake actions together. Given that shared agency is the primary way by which we shape and navigate our social world, we develop models of shared agency, in part, to manage how we engage in this practice. These models are thus a second-order form of social reality management because their purpose is to manage the conditions by which we create social reality together. For this reason, we ought to be concerned with both the ideological nature of these models and the ethics of this second-order management. Contra universal or singular models, I argue that shared agency manifests differently across conceptual schemas and through non-ideal conditions, drawing in particular on decolonial feminist theory. Thus, we ought to strive for a multiplicity of models, particularly in order to narrow (or close) the gap between a model and those whose agency is managed by that model. By starting from localized, bottom-up models, research on shared agency also operates as a site of radical possibility allowing us to democratically explore what new forms of shared agency could be.

TABLE OF CONTENTS

Chapter	Page
INTRODUCTION	1
Chapter Abstracts and Overview	6
CHAPTER I	26
METHODOLOGY AND IDEALIZATIONS IN SHARED AGENCY RESEARCH	
Section I. The ‘Social’ in Social Ontology	26
Section II. Hans Vaihinger: Thought, Purpose, Models	32
Section III. Idealizations, Ideal Theory, and Oppression	38
Section IV. Concluding Remarks	49
CHAPTER II	50
SHARED INTENTIONS AND IDEALIZATIONS	
Section I. The Question of Methodology	50
Section II. Shared Intentions	53
<i>i.</i> Michael Bratman: Shared Planning Agency	54
<i>ii.</i> Margaret Gilbert: Joint Commitments	61
<i>ii.</i> Christopher Kutz: Participatory Intentions	68
Section III. The Methodology for Shared Agency	74
<i>i.</i> Bratman: Creature Construction	79
<i>ii.</i> Gilbert: Personal Readiness	82
<i>iii.</i> Kutz: Rational Choice Theory	85
Section IV. General Analysis	87

CHAPTER III	90
SHARED INTENTIONS AND UNJUSTIFIED IDEALIZATIONS	
Section I: The Purpose of a Model	90
Section II: Structural Oppression and Shared Agency	93
Section III: Three Models and Their Idealizations	109
<i>i.</i> Idealized Social Ontology	110
<i>ii.</i> Idealization of Agent Neutrality	115
<i>ii.</i> Idealized Epistemology	121
Section IV: Ideology and Idealizations	131
CHAPTER IV	134
MODELING SHARED RESISTANCE	
Section I: Whose Purpose? Whose Values?	134
Section II: Epistemology and Intersectionality	137
Section III: Case Studies	153
<i>Case 1:</i> Building a Community Together	153
<i>Case 2:</i> Opposing the _____ Together	164
Section IV: Concluding Remarks	172
CHAPTER V	173
MODEL PLURALISM AND THE ETHICS OF SOCIAL REALITY MANAGEMENT	
Section I: Introduction	173
Section II: Open-Ended and Constrained Model Pluralism	176
Section III: Constrained Model Pluralism	179
Section IV: Management and Justification	182

Section V: Closing the Gaps Between Manager and Managed	193
Section VI: Friction and Forms of Life	205
CONCLUSION	219
REFERENCES	233

Introduction

Social reality management refers to the myriad ways through which we construct, reshape, and otherwise organize features of our social worlds, including ourselves as social beings, in accordance with the various needs and values of human life. My choice of the word ‘management’ is intentional over other options, such as engineering, navigation, stewardship, or governance, but it is ultimately how unsatisfactory the word itself is that constitutes my provisional usage of it at this stage.¹ Management can take many different forms and for this reason tracks an array of both negative and positive connotations. Social engineering invokes images of eugenics, even if it ought to equally bring forth images of mass resistance to eugenics, invoking also a macro-level sense of telos that I want to avoid. Navigation suggests that we are all equal in this process and does not appropriately invoke the notion that navigation itself changes that which is being navigated. Stewardship and governance also fail because they have a top-down connotation, even if their actual practices do not take this form.

Management, however, does not avoid similar connotations but rather seems to include them all. As opposed to maintenance, it invokes some sense of telos in virtue of our particular purposes for management, but it need not be macro-level nor singular. One can manage their garden or a forest or engage in communal management as a member of the community, but one can also be managed in justified or unjustified ways. Parents manage their children, for better or for worse, but a failure to manage one’s child at all would be neglect. Certain forms of law manage, for better or worse, the relationship between the citizens, each other, and the state.

¹ To be clear, I am adopting the word used by two central theorists in my dissertation, Hans Vaihinger and Kwame Anthony Appiah, but I want to be clear that I am intentionally adopting it and critically so. That is, this word is ultimately unsatisfactory, but at the present, there is not another word that better captures what I am describing. My use here is provisional in that sense.

Driving laws are a helpful form of social reality management (even if a privatized and individualized framework for transportation is not), but also can be constructed and enforced in deeply unjust ways. Moreover, one can manage in many directions. Governance is often, although need not necessarily be, a macro-level, top-down form of management, but grassroots forms of social justice organizing constitute a form of micro-level, bottom-up management.

Management thus can capture an open-ended set of actions related to how people construct, reshape, and organize social reality without emphasis on a single or macro-level goal nor excluding concern for unjust management. That is, the term itself has a kind of uneasiness to it, which is why I've chosen it. Social reality management can be oppressive, shaping our social-ontological landscapes and our positions within them in unjust ways. The sense of being managed is often unpleasant and invokes the image of a corporate-like structure of domination in which the structure itself is already the outcome of previous forms of management. People tend to not like their managers, especially if they micro-manage, nor unsolicited advice about how they ought to manage their own life. Management, of course, can also be empowering, functioning in many cases as a form of resistance to mismanagement. The ability to self-manage one's work can be liberatory, but, of course, the ability to self-manage or equally co-manage is precisely what forms of unjust social reality management deny. For this reason, the uneasiness of the term serves as a reminder that shaping and organizing social reality can have morally and politically horrific consequences (i.e., the case of eugenics and genocide) if not done critically and with care, that it can drastically alter our social worlds and relations to others (i.e., forms of social categorization such as gender, race, and nationality), that the scope of management itself requires further justification as one cannot be a social reality tyrant, and that who is managing (or

who is denied the ability to manage),² what they are managing, and why they are justified in doing so is the most central concern of this process. Our social worlds do not manage themselves, and pretending otherwise simply obscures widespread forms of mismanagement by those who directly benefit from it.

Social reality management is also a shared process. Thus, my focus for this dissertation is on how people manage the terms of this shared process: the process of sharing agency. I argue that the philosophical study of shared agency is about the management of how people shared agency. In the case of analytic literature on shared agency, theorists often offer highly general models for how two or more people successfully share agency. These models are normative, serving as an ideal for how people ought to engage in this process successfully. The application of such models thus has the power to regulate this process, determining what counts as a genuine form of shared agency, the terms of success or failure, and offering a broader normative claim about what agency is, shaping how agents come to think of themselves. In this way, providing a model for shared agency is a second-order form of social reality management insofar as it manages the terms by which we undertake this process together and who is regarded as the right kind of agent to engage in this process. My dissertation offers a methodological, ideological, and ethical analysis of three predominate models within this literature, arguing that their ideal methodology and particular use of idealizations results in models that are unjustified forms of

² There is also a corollary problem of people unequally tasked with certain forms of management, such as management of the household or management of institutional diversity, inclusion, and equity programs. In this way, being tasked with managing can also be a way to deny people the ability (or time) to manage other things. I raise this as an important concern about oppression and social reality management but do not pursue this here. For further discussion in the case of microaggressions, see: Christina Friedlaender, "Microaggressions and the Problem of Attributional Ambiguity," in *The Routledge Handbook to Social and Political Philosophy of Language*, eds. Justin Khoo and Rachel Sterken (New York: Routledge University Press, *forthcoming*).

second-order management.³ My focus on three different models is to ultimately highlight a broader methodological trend in the literature – one that requires critical attention.

My critical project ultimately makes a distinction between how theorists of shared agency ought to proceed and how they have proceeded so far. First, if our models for shared agency are not power-sensitive, tracking micro- and macro-level forms of oppression, then they necessarily risk being unjust in their own right. In particular, I argue that an ideal methodology can offer, at best, a *constrained model pluralism*. Constrained model pluralism is the assumption that a universal or limited set of highly general models is sufficient for understanding and managing shared agency. If a model aims to capture how large-scale forms of shared agency work, then it does so by offering a highly general model, rather than developing a large-scale model for a particular purpose through first attending to differences in localized models between those who see a need for the large-scale model in the first place. It is also constrained in the sense that claims to value pluralism are often constrained by appeal to parochial standards of reasonableness. Because the idealized models within the current literature are normative, they appeal to normative accounts of agency in which norms of rationality are universal and third parties can assess the reasonableness of other agents in universal terms as well. This necessarily excludes all values that do not meet such standards, and similarly, constrained model pluralism excludes all forms of shared agency that do not match the value structure of an ideal methodology, which is reflected particularly in the kinds of idealizations such theorists employ. Lastly, constrained model pluralism offers static models for what shared agency is or should be, treating a dynamic and heterogenous social phenomenon as static and homogenous.

³ My arguments about ideal theory and idealizations in the case of shared agency pulled from my initial discussion here: Christina Friedlaender, “Resisting Oppression Together: Shared Intentions and Unequal Agents,” in *Collectivity: Ontology, Ethics, and Social Justice*, eds. Kendy Hess, Violetta Igneski, and Tracy Isaacs, 265-289 (London: Rowman & Littlefield International, 2018).

Open-ended model pluralism, on the other hand, avoids constraints of this sort. The claim here is that both understanding and managing the process of shared agency works best by starting from particular, localized models. Such models are never fully complete, but rather arise from and are tested through forms of social action. That is, modeling shared agency is itself a shared process. A multiplicity of models, moreover, allows for models that more accurately reflect global and historical diversity in metaphysical worldviews, values and value structures, and social practices (instead of starting from a parochial metaphysical worldview under the guise of universality and neutrality). Since social reality management will often require the ability to share agency on a large-scale level, there is still a need to develop models through which those who do not significantly share social backgrounds can come to genuinely share agency toward a particular goal. Thus, I argue that open-ended pluralism places these partial models into friction with one another, allowing people to derive a large-scale model in virtue of the differences and similarities between them.

Open-ended model pluralism pushes us to consider not only what forms of shared agency already exist but also opens up an imaginative space for new possibilities of this social phenomenon. In this way, theorizing shared agency as a form of second-order social reality management is ultimately about future possibilities for our social-ontological landscapes. We do not all imagine the future in the same way. Some visions of the future exclude entire groups of people and ways of life, although exclusion itself is not necessarily unjustified. Some visions are vaguer than others because oppressive socio-ontological landscapes can deny people the ability to imagine their future, both individually and collectively; and thus, how we develop partial models for shared agency, for what purpose we develop them, and what values or metaphysical assumptions they include comes down to a question of what kind of futures we are planning

together, to what extent our various visions are or can be shared, and who is a part of this planning and thus a part of the future, be it in determining the future of a marriage, the future of a nation-state, or the future of earth as a habitable planet and for whom it will be habitable. Thus, second-order social reality management must necessarily address histories of exclusion and oppression that inform our present and often unjust socio-ontological landscapes, because such landscapes inform our possibilities for the shared agency and planning for the future together.

Chapter Abstracts and Overview

Chapter I: Methodology and Idealizations in Shared Agency Research

Over the past fifty years, there has been much debate about the usefulness of ideal theory as a methodology and its method of idealization. The initial division between ideal and non-ideal theory comes from John Rawls' *Theory of Justice*.⁴ An ideal theory is a methodological process for first arriving at a universal set of moral, political, or otherwise normative principles that can then be applied to particular, non-ideal contexts. Broadly construed, idealization is any form of abstraction, generalization, or simplification that introduces a false fact into a model and, in doing, so aids in the use of the model for the purposes of understanding, predicating, or otherwise managing the phenomenon under investigation. Ideal theorists employ the method of idealization by beginning from a set of favorable assumptions about agents, their capacities, and the social-ontological context(s) in which such agents exist. For Rawls and others, an ideal methodology functions as a form of justification for the set of normative claims, specifically the universality of such claims. This famously appears in Rawls' *original position*, wherein agents,

⁴ John Rawls, *A Theory of Justice, revised edition* (Cambridge, MA: Belknap Press of Harvard University Press, 1999).

who have no substantive identities, interests, or personalities, and who lack any particular knowledge of the social, political, historical, or geographical circumstances in which they exist, will also arrive at the same two principles of justice.⁵ Rawls derives his ideal methodology by appealing to Immanuel Kant's conception of moral agency, specifically that all rational beings will arrive at the same universal moral principles in taking the standpoint of reason.⁶ That is, the methodology of ideal theory is necessarily intertwined with a particular normative conception of agency as well as an appeal to universality.

In Chapter I, I begin the first claim of my core argument that much of the analytic literature on collective intentionality relies on an ideal methodology and employs a similar set of idealizations. Like their moral and political counterparts, many of the accounts within the literature provide normative models for the success conditions of collective intentionality. The main difference is that unlike the case of moral and political philosophy, this subfield of social philosophy largely aims to develop models that have a highly general scope of application, but a few do, in fact, take up universalizability as their aim. Nevertheless, the level of generality at which these models aim is often nearly universal in scope. My particular focus is on the case of shared intentionality; that is, how two or more people share an intention to undertake an action together. I examine three of the most predominate accounts in the literature: Michael Bratman's shared planning agency, Margaret Gilbert's joint commitment, and Christopher Kutz's participatory intentions.⁷ Between Chapters I-III, I argue that these three accounts employ a similar set of unjustified idealizations – (1) an idealized social ontology, (2) the idealization of

⁵ Rawls, *A Theory of Justice*, 15-19.

⁶ Rawls, *A Theory of Justice*, 221-227.

⁷ Michael Bratman, *Shared Agency: A Planning Theory of Acting Together* (Oxford: Oxford University Press, 2013); Margaret Gilbert, *Joint Commitment: How We Make the Social World* (Oxford: Oxford University Press, 2014); Christopher Kutz, *Complicity: Ethics and Law for a Collective Age* (Oxford: Oxford University Press, 2000).

agent-neutrality, and (2) an idealized epistemology – that result in methodologically inferior models for how two or more people successfully share an intention, which are also ethically and ideologically concerning.

Thus, Chapter I covers the first part of this argument by examining the relationship between methodology and the method of idealization in how theorists develop a model for some feature of social reality – in this case, shared intentions, or, more broadly, shared agency. Drawing on Hans Vaihinger’s *The Philosophy of ‘As If’’: A System of the Theoretical, Practical, and Religious Fictions of Mankind*,⁸ I examine why it is we create models of various phenomena within reality and what role idealizations play within the creation of such models. Vaihinger argues that we model features of reality for two related reasons: (1) to enhance our understanding of the phenomena under investigation; and (2) to use this enhanced understanding to manage the phenomena in question for the purposes, needs, and values of human life. The relationship between understanding and management is clear in the case of normative models, but it holds true for descriptive models as well. Both descriptive and normative models are created for a particular purpose, and given that this purpose is always in relation to the needs of human life, both forms necessarily contain evaluative commitments or expressions of value about the phenomenon under investigation.

Following Vaihinger’s fictionalism (or proto-pragmatism), I examine the debate between ideal theorists and oppression theorists, a particular subset of non-ideal theorists who reject ideal theory *as well as* the use of idealizations in normative theorizing. I argue that oppression theorists are methodologically wrong to eschew the use of idealizations, pointing out that such theorists do not actually avoid them in practice. More broadly, however, I show that ideal and

⁸ Hans Vaihinger, *The Philosophy of ‘As If’’: A System of the Theoretical, Practical, and Religious Fictions of Mankind*, trans. C.K. Ogden (New York: Harcourt, Brace, and Co., 1924.)

oppression theorists use the method of idealization in drastically different ways. Ideal theorists idealize *away* from the complexities of co-forming systems of power and domination in order to arrive at a normative account that posits an ideal by which we can, in our non-ideal application, determine how non-ideal conditions fail to match and how to increase that match.⁹ Oppression theorists, on the other hand, idealize *from* the concrete conditions of particular socio-political and historical contexts, providing normative accounts for the betterment of those particular non-ideal conditions but without a static ideal as the guide. Thus, in this chapter, I argue that idealizations are methodologically necessary, even for oppression theorists, such that my criticism about the predominate models for shared agency – borne from oppression theory – is about their particular use of idealizations. Looking forward, my solution to this problem rests on the contextually justified idealizations of oppression theory, not eschewing idealizations altogether.

Chapter II: Shared Intentions and Idealizations

Chapter II serves two purposes in the context of my dissertation. First, I provide an overview of three predominate models of shared agency within the analytic literature. Since my underlying claim is that an ideal methodology is present throughout much of the current literature, I focus on three dissimilar models that capture the broader contours of the literature. Second, I demonstrate how each model relies on a similar and related set of idealizations: (1) an idealized social ontology; (2) the idealization of agent-neutrality; and (3) an idealized epistemology. Third, drawing on Kwame Anthony Appiah's recent analysis of Hans Vaihinger's methodology,¹⁰ I begin my argument that serves as the core of Chapter III: these three models all

⁹ Charles W. Mills, "'Ideal Theory' as Ideology," *Hypatia* 20, no. 3 (Summer 2005): 168. Mills uses the particular phrase of 'abstract *away*,' which I am directly borrowing here.

¹⁰ Kwame Anthony Appiah, *As If: Idealization and Ideals* (Cambridge, MA: Harvard University Press, 2017).

employ an ideal methodology, which is a problem in itself, and that the standard forms of idealization are contextually unjustified, or, more simply put, bad. This critique is not merely about methodology, but also concerns the ideological nature of an ideal methodology and these idealizations.

My overview focuses on three different models. First, Michael Bratman's model for shared planning agency offers a set of five success conditions for how two or more people successfully share an intention together.¹¹ This takes the form of: I intend that we *J.* Bratman's model relies on a process of creature construction and augmented individualism, namely that large-scale social phenomena can be explained by way of building together small-scale social phenomena.¹² That is, shared planning agency can be explained by slightly modifying individual planning agency, remaining ontologically parsimonious by not postulating new intentional states or forms of agency. Bratman's focus is on small-scale actions that can easily be carried out by a dyad or triad, such as painting a house together. His focus on simple cases also excludes all forms of hierarchy, authority, or power – be it formal or informal, justified or not – as well as the presence of a broader social and political world in which this small-scale action occurs.¹³ The agents in question lack any substantive features, except for the capacity for practical rationality and the ability to engage in rationality-structured, future-oriented planning. Finally, Bratman argues that his small-scale account of shared planning agency is further scalable, offering a broader picture of how large-scale forms of shared planning occur by focusing specifically on the terms of shared deliberation.

¹¹ Bratman, *Shared Agency*, 152.

¹² Bratman, *Shared Agency*, 11-15.

¹³ Bratman, *Shared Agency*, 7.

Margaret Gilbert's model for joint commitment breaks substantially from Bratman's parsimonious model by postulating both a new intentional state and kind of agency. Her model for joint commitment takes the form of: we intend to *J*. A joint commitment, unlike a personal commitment to one's self, is a commitment between two or more people to φ as a body.¹⁴ This means that a joint commitment creates a *plural subject*,¹⁵ which is not the amalgamation of individual wills but rather an indivisible subject in its own right. A joint commitment is not necessarily an agreement, but rather arises when all participating parties express a *personal readiness* (i.e., an expression of willingness) to enter into the commitment and thereby become a plural subject.¹⁶ Gilbert's model is more explicitly normative than Bratman's, claiming that joint commitments are defined in virtue of their answerability function. If one member decided to unilaterally rescind or break the commitment, then the other parties are justified in normative rebuke. Unlike Bratman, Gilbert does not exclude the presence of power in relations between subject, nor does her account provide a nuanced analysis of how forms of power impact joint commitments. She does, however, argue that individual agents must be statistically normal humans in terms of perceptual and conceptual capacities. Lastly, while Gilbert similarly starts with small-scale actions, such as taking a walk together, she argues that joint commitments are not merely scalable but, in fact, form the structure of our social and political worlds, claiming that societies or nation-states can themselves be plural subjects.¹⁷ Gilbert's model for joint commitment thus aims to be universalizable, whereas as Bratman's aims to be highly general.

¹⁴ Gilbert, *Joint Commitment*, 7.

¹⁵ Gilbert, *Joint Commitment*, 8-9.

¹⁶ Gilbert, *Joint Commitment*, 47-48.

¹⁷ Margaret Gilbert, *A Theory of Political Obligation* (Oxford: Oxford University Press, 2006).

Finally, Christopher Kutz's model for participatory intentions is more ontologically parsimonious than either of the above, removing the pronoun 'we' altogether. A participatory intention is an individual intention to contribute to a collective outcome, even if the individual does not intend the realization of that end.¹⁸ Kutz's model for participatory intentions is built on a pro-social game theory in which participants might have knowledge of others' preferences for action but also can determine or revise their preferences in light of others. In this way, Kutz's model is not a standard form of shared intention, focusing instead on an individual's phenomenological sense of agency within group-oriented contexts (i.e., the sense of acting in relation or in tune with others toward a collective outcome).¹⁹ One can carry out a participatory intention alone in the hope that others will, upon witnessing this action, also develop participatory intentions, rendering the action a collective one. Unlike the more substantive common knowledge conditions required by Bratman and Gilbert, Kutz's success conditions offer a lower threshold of *mutual openness*, which does, however, require agents to have a significantly shared conceptual background.²⁰ Lastly, Kutz's model aims to track all kinds of collective actions: small or large, planned or spontaneous, and tracks forms of hierarchy, be it formal or otherwise.²¹ Because Kutz's model is tied to broader moral questions surrounding complicity, his model for participatory intentions aims to be universal.

Bratman, Gilbert, and Kutz each provide a model for shared agency that follows an ideal methodology in which an ideal model is applied to non-ideal conditions, using the three standard idealizations in their own ways. First, all three start from an idealized social ontology in which

¹⁸ Kutz, *Complicity*, 94.

¹⁹ Kutz, *Complicity*, 5.

²⁰ Kutz, *Complicity*, 76-77.

²¹ Christopher Kutz, "Acting Together," *Philosophy and Phenomenological Research* 61, no.1 (2000): 9.

forms of social inequality and structural oppression do not exist,²² and thus have no bearing on: (1) the development of agents and their sense of agency; (2) claims about who constitutes an agent and who is an appropriate agent with whom one can attempt to share agency; (3) the role that inter-connected systems of power have on the forms and possibilities of shared agency; and (4) the social meaning of the actions themselves and in relation to the agents undertaking those actions, among other items.

Second, each relies on the idealization of agent-neutrality. Agent-neutrality provides a fictional picture of agents in which their social, political, cultural, and religious social worlds are taken to be contingent rather than partially constitutive features of their agency, such that rationality and reasonableness are treated as the necessary and relevant features of agency for the models. Like Rawls' original position, agents are essentialized to a few capacities, largely practical rationality and the ability to interpret the intentional states of others with a relative ease. This idealization relies on a two-fold assumption. First, human agency can easily be divided in terms of essential and contingent features, because the so-called contingent features have no direct bearing on how agents understand, interpret, and act in relation to others. Second, if there are a set of essential features, everyone ranks the priority of these features similarly or understand these features to similarly to others. For example, autonomy is taken to be one important feature of agency, but the evaluative ranking of this feature is not globally shared, nor is autonomy understood in the same terms (i.e., individual versus communal autonomy).

Lastly, each account relies on an idealized epistemological framework. To say an epistemological framework is idealized is that it is fundamentally asocial and apolitical, assuming that agents have a largely shared set of hermeneutical resources for interpretation of

²² Mills, "'Ideal Theory' as Ideology," 168.

each other and the nature of the action they are undertaking. This idealization also includes, what Charles Mills calls, an idealized cognitive sphere which presupposes there is a “general social transparency.”²³ Their epistemological framework is asocial insofar as knowers are not taken to be situated within a particular socio-ontological landscape, and therefore knowledge is not produced from a particular location (and thus reflective of that location in some capacity), but rather a form of universal discovery accessible to all agents in feature of their essential capacities. The use of these three idealizations thus broadly removes the presence of forms of structural oppression and the relationship between such structures, agency, and knowledge.

Chapter III: Shared Intentions and Unjustified Idealizations

Chapter III builds on the central claims of Chapter I and II, arguing that an ideal methodology and the three kinds of idealizations used within models for shared agency result in methodologically inferior and ideologically concerning models. In Chapter I and II, I argue for three central claims: (1) we model phenomena, particularly social phenomena, in order to increase our understanding and enhance our capacity to manage the phenomena in accordance with the various needs and values within human life; (2) idealizations are a necessary feature of such models, but oppression theorists idealize differently from ideal theorists; and (3) Bratman’s, Gilbert’s, and Kutz’s models for shared agency all use three standard idealizations, albeit in their own ways. Chapter III builds on these claims by arguing that there is an intimate relationship between how we model the social world in order to manage it and how we idealize from our dynamic, power-infused social world that is always being constructed and reconstructed in virtue of our attempts to manage it.

²³ Mills, “‘Ideal Theory’ as Ideology,” 169.

The first part of the chapter details the ways in which oppression theorists idealize *from* the realities of structural oppression, starting from the non-ideal world rather than reintroducing such conditions into an ideal model during the stages of application. I note three forms of non-ideal idealizations. First, the use of general hyponyms allows theorists to refer to highly complex, interconnected structural forms of oppression, tracking the interplay between micro- and macro-systems of power, and what bearing this has on the development of agency, particularly its formations in relation to forms of and shifts within social categorization. Such hyponyms are also sites of critique for how such idealizations can be useful to motivate certain kinds of structural critique but also, if used poorly, also result in forms of over-generalization. Thus, the use of hyponyms is a critical process in which theorists are able to idealize from micro-processes of power to establish claims about macro-processes, while, at the same time, ensuring that the macro-claims remain necessarily in a dialogue with micro-claims. This way, models reflect a non-idealized social ontology. Second, while an ideal methodology relies on the idealization of agent-neutrality, oppression theorists offer ways to theorize about similarity and difference in the case of identity without essentializing such forms of social categorization and construction. This is to say that identities themselves constitute useful idealizations that allow one to attend to how the construction and internalization of forms of social categorization impact possibilities for shared agency, particularly between agents who have different *social-structural positionings* within micro- and macro-systems of power.²⁴ Lastly, drawing in particular on an intersectional methodology, the use of metaphor (as well as related heuristics) itself is a powerful form of idealization, because metaphors can radically shift how theorists envision the contours of

²⁴ The phrase ‘social-structural position’ comes from: Iris Marion Young, *Responsibility for Justice* (Oxford: Oxford University Press, 2011), 19, 45.

oppression within the process of theorizing.²⁵ Metaphors offer a particular kind of fiction, one that can offer a more precise starting point for imagining the dynamic co-formation and interplay between micro- and macro-systems of power, the relationship between agency, agents, and these complex processes, and, in particular, the power-laden relationship between knowledge production, experience, and resistance. Metaphors are more precise than non-metaphorical language because they can radically shift how we perceive our social-ontological landscapes, our own locations within such landscapes, what kinds of things we take to be evidence and our perception of such evidence, and, by extension, how to theorize about our social world.

In these three ways, oppression theorists construct more localized models for forms of shared agency that track an unjust and oppressive social ontology, agential differences, and the social and political dimensions of knowledge. Contra an ideal methodology and the three standard idealizations, this oppression-focused approach starts from concrete, non-ideal conditions, idealizing in order to capture our highly complex and dynamic social worlds (as opposed to removing complexity altogether until the stage of application). Following Vaihinger, these idealizations are contextually justified, or good, because they track a more narrowly tailored purpose of a more narrowly tailored model. In this way, oppression theory uses idealizations but does so in ways that allow people to better understand and manage complex systems of power (wherein management is often conceptualized in terms of resistance).

Lastly, my claim is not simply that the three idealized models for shared agency are inferior for the process of understanding and management, but rather that such models are ideologically concerning insofar as their methodological procedure and idealizations naturalize processes of social construction and ignore how micro- and macro-level unjust structural

²⁵ Patricia Hill Collins, *Intersectionality as Critical Social Theory* (Durham, NC: Duke University Press, 2019), 29-45.

processes impact agents. I argue that if our models for shared agency aim to manage how it is we actually engage in this practice, then idealized models cannot (and do not) critically assess whether and to what extent their models risk adopting forms of unjust management (or the outcomes thereof). In this sense, as Charles Mills argues, ideal theory and its particular use of idealizations is concerningly ideological because the decision to exclude concerns about power already reflects the interests of those who largely benefit from oppressive forms of social reality management.²⁶ However unintentional, such models thus risk adopting oppressive social-ontological landscapes, having naturalized them. This kind of adoption in a model used for management will reiterate such forms of oppression through its use. Because an ideal methodology purports to offer an impartial or objective approach to normative projects, it uses this claim to impartiality as evidence that it is ideology-free, that those who are critical of supposedly neutral or impartial starting points are themselves the ones wrapped up in ideology.

Finally, given my claim that Bratman's, Gilbert's, and Kutz's models cannot merely reintroduce non-ideal complexity within the stage of application, I analyze how each particular model fails in virtue of the idealizations it uses by following through on their methodological process and reintroducing non-ideal complexity. In the case of Gilbert and Kutz, this is relatively easier to show because, despite the universal nature of these models, their particular success conditions around common knowledge or mutual understanding necessarily require a significantly shared conceptual and perceptual background between agents. Because Bratman does not offer a universal model, I highlight how his exclusion of power ignores how it can be constitutive of both agency and relations between agents, even in the small-scale actions of a dyad. Here, my ideological claim about these models raises an ethical concern about how

²⁶ Mills, "'Ideal Theory' as Ideology."

theorists model shared agency. If the reason we model forms of shared agency is ultimately about management, then we ought to be ethically concerned with the nature of this management, particularly in terms of who is justified in managing what, to what extent, and why. Thus, Chapters IV and V take up my ideological claim and ethical concern more deeply, offering an alternative approach to the process of modeling itself.

Chapter IV: Modeling Shared Resistance

Chapter IV further develops my methodological and ideological argument about reliance on an ideal methodology. Connecting Hans Vaihinger's claims about why we create models of reality to the traditions of feminist epistemology and feminist philosophy of science, I show how shared agency models necessarily reflect parochial values and ontological assumptions, which is not bad in and of itself, but does mean that such models cannot be easily applied contexts that do not reflect these values. However, failure to understand a model's limitations (and why those limitations exist) in this sense often reflects a social and political epistemic failure. This chapter offers two case studies that: (1) further highlight why idealized models cannot simply re-introduce non-ideal conditions at the stage of application; and (2) how knowers, and thus knowledge, are situated, such that there is an intimate connection between power, experience, and ignorance.

My analysis of such cases starts first in a broader discussion of non-idealized epistemology. Following standpoint theory, I emphasize that knowers are situated within the world, such that their knowledge claims often reflect features of their situatedness. As with Vaihinger, standpoint theory is broadly the claim that our forms of knowledge production reflect our particular needs and values as humans, that (1) we do not all share the same set of needs and

values, and (2) situated knowledge is partial and dynamic.²⁷ Unlike Vaihinger, standpoint theory is more particularly a claim that forms of structural oppression and other expressions of power impact our situated knowledge projects. As part of this analysis, I look at how an intersectional methodology and set of epistemological practices offers a deeper, more complex framework for standpoint theory,²⁸ which counters its historically monolithic approach to the category of ‘woman.’ An intersectional methodology rejects the three standard idealizations (i.e., agent-neutrality, idealized social ontology, and idealized epistemology) at every stage in the process of theorizing, by arguing that a multi-axis approach to theorizing about oppression and social categorization requires anti-segregation and anti-subordination thinking.²⁹ Moreover, this approach frames theorizing itself as a shared practice that arises within the particular contours of an interpretative community.³⁰ An interpretative community has its own set of practices for knowledge production, including legitimation and justification. In this case, resistant knowledge projects, similar to group standpoints, are the process by which members of oppressed groups come to understand the terms of their oppression for reasons of survival and flourishing.³¹

²⁷ See, e.g., Sandra Harding, *Whose Science? Whose Knowledge?: Thinking from Women’s Lives* (Ithaca, NY: Cornell University Press, 1991).

²⁸ This is not to argue intersectionality, as both a set of epistemological practices and as a resistant knowledge project in its own right, is the same as standpoint theory. On the one hand, intersectionality offers a richer account of how to develop group standpoints, the limitations of a standpoint, and how standpoints relate to social group categorization within ontological forms of oppression. However, it also offers a way to think about individual knowledge claims (not just group-level ones) and its focus on interpretative communities helps shift the social-group-centric approach of standpoint to a broader claim about differences in micro- and macro-level communities. I take this shift to community to be more valuable because the hyper-focus on social groups sometimes begets a single-axis or over-idealized multi-axis approach to knowledge production and appeals to standpoints, often treating forms of social categorization as static and universal (and, by extension, knowledge claims as such as well). See: Collins, *Intersectionality as Critical Social Theory*, 116-120, 136-142

²⁹ Kimberlé Crenshaw, “Demarginalizing the Intersection of Race and Sex,” *University of Chicago Legal Forum*, 140, no. 1 (1989): 139-167. See also: Anna Carastathis, “Basements and Intersections,” *Hypatia* 28, no. 4 (Fall 2013): 698-715.

³⁰ Collins, *Intersectionality as Critical Social Theory*, 127-136.

³¹ Collins, *Intersectionality as Critical Social Theory*, 127-136.

Thus, what this analysis highlights in the case of shared agency is that a supposedly neutral or impartial approach to developing a model fails to appropriately grapple with embedded values. The three idealizations currently used cannot account for conceptual schemas in which we are embedded, particularly in organizing within intertwined micro- and macro-systems of power and oppression. Such idealized models thus cannot provide an enhanced understanding of how particular attempts to shared agency fail in virtue of assumptions of universality, leaving users unable to address forms of joint action that require participants (or potential participants) to directly address these differences in a power-sensitive way. Because such models cannot accommodate these concerns in virtue of their appeals to universality and neutrality, the function of such models not only risks reinforcing oppressive practices, but, in many cases, a model for shared agency can necessarily reinforce oppression in virtue of the exclusionary terms and socio-ontological landscapes that allowed that model to develop in the first place.

Finally, I offer two case studies in this chapter in order to highlight the failure of these idealized models, particularly by arguing that taking on a power-sensitive and justice-oriented approach is not sufficient to avoid the problems of these models. That is, I focus on cases in which participants (in one way or another) take themselves to be engaging in either a form of anti-oppression resistance and/or contending with the failures within others' attempts at resistance that directly or indirectly affect them. Case 1 follows the organization of a *Signs* conference on the topic of 'communities of women' at Stanford University in 1983. My goal is to show 'success' and 'failure' are interpreted differently among the organizers, participants, and potential participants, particularly in relation to the organizers' narrow and exclusive model for what 'community' is, and how this functions in the case of feminism. The second case study

focuses more on the importance of shared concepts and terminology for shared intentions and joint actions. I examine the recent controversy surrounding Alexandria Ocasio-Cortez's use of the term 'concentration camp' to refer to facilities run by the United States Immigration and Customs Enforcement agency. By examining the difference between 'concentration camps' and 'detention facilities,' I show how conceptual and linguistic differences can alter shared intentional context, both in terms of who is a participant as well as the meaning and structure of the action itself. Both cases offer a direct counter to an idealized epistemological approach, demonstrating that the contours of epistemic difference (in practices and in interpretation) require deeper attention to the relationship between epistemic insight, social-structural positionings, and structural forms of ignorance that impact people differently in virtue of their positionings. In this way, each study captures a different set of problems with these models. My point here is to show that due to their idealizations, the three idealized models cannot adequately address these cases, and, even worse, the use of these models to analyze these cases would label certain forms of failure as success.

Chapter V: Model Pluralism and The Ethics of Social Reality Management

Chapter V brings together my methodological, ideological, and ethical critique of an ideal methodology and its particular idealizations in the case of shared agency. I argue that enhancing both our understanding of shared agency and our ability to manage it in its various forms requires a methodological commitment to *open-ended model pluralism*. Moreover, I argue that the use of a universal model or a limited set of highly general models constitutes a tyrannical form of second-order social reality management. Following Chapter IV, I argue that open-ended model pluralism is better for reducing the gap between the model, the creators of the model, and those who are managed by such models. Self-management is a case in which there is no gap,

whereas certain forms of communal autonomy constitute a minor gap in comparison with an externally imposed and unwelcome model. That is, open-ended model pluralism allows theorists to directly address the ethical relationship between the managers and the managed at both micro- and macro-levels.

Drawing on the distinction between value pluralism and reasonable value pluralism often found in moral and political ideal theory, I contrast open-ended model pluralism with *constrained model pluralism*. The difference between the two broad approaches is fundamentally about: (1) how localized and how many models we ought to strive for; (2) how we ought to approach developing models for larger-scale forms of shared agency; (3) the allowance for new possibilities for forms of shared agency; and (4) how power-sensitive our approach needs to be. A constrained approach, like ideal theory, focuses on a more limited set of highly general models in which there is a predetermined claim about what models might be reasonable in virtue of how we assess agential rationality, particularly in universal terms. Open-ended model pluralism, on the other hand, shifts the focus to how partial models for shared agency develop through actions within and between various communities, offering an approach that treats theorizing about shared agency as a fundamentally shared practice in itself that occurs within and between different interpretative communities.

Because an ideal methodology is ideologically connected to both comprehensive and political liberalism, I further highlight the ethically concerning nature of such models in terms of social reality management by focusing on discussion of the relationship between missionary, or colonial, feminism, and forms of decolonial feminism.³² Attempts to shared agency in the case of transnational feminist projects illuminate what happens when even non-idealized models fail to

³² In particular: Serene Khader, *Decolonizing Universalism: A Transnational Feminist Ethic* (Oxford: Oxford University Press, 2019).

critically grapple with contextual differences between socio-ontological landscapes of oppression, rendering them similarly criticizable to idealized models. In particular, I show how constrained model pluralism results in a tyrannical form of second-order reality management, showing how this happens in the case of colonial feminism.

Finally, connecting open-ended model pluralism to Ludwig Wittgenstein's discussion of language games and forms of life, I argue that investigation into shared agency should not be a top-down form of conceptual refinement, because shared agency operates differently in different language-games, tracking, more broadly, different forms of human life.³³ That is, there is no single answer to what shared agency is. Theorizing through partial and dynamic models for shared agency needs to focus on what possibilities for shared agency arise from our current socio-ontological landscapes, offering a better way to approach the need for models can be used by communities who do not significantly share social world(s) (but whose social worlds likely developed within some relation to one another in more direct or indirect ways).

Using José Medina's work on epistemic resistance,³⁴ I explain how this open-ended pluralism allows for the possibility of *epistemic friction* between models and thus encourages *the interaction of resistant imaginations* between forms of shared agency. In the case of resistance, a greater plurality of partial models is more useful, especially when such models develop through reflection on the experiences of those who actively participate in their development. In this sense, open-ended pluralism opens up possibilities for the genuine democratization of both second- and first-order social reality management. If the goal is to develop ways to massively share agency between people who may be operating off of many different and incompatible

³³ Ludwig Wittgenstein, *Philosophical Investigations*, trans. G.E.M. Anscombe (Oxford: Basil Blackwell, 1958).

³⁴ José Medina, *The Epistemology of Resistance: Gender and Racial Oppression, Epistemic Injustice, and Resistant Imaginations* (Oxford: Oxford University Press, 2013).

models of shared agency³⁵ – for example, in the case of large-scale forms of resistance or revolution – then it is the friction between these models that opens up a space to imagine the possibilities for massively shared agency in the first place. That is, it is only through a multiplicity of more localized models that we can even come to develop more general models in order to share agency on a larger scale. In this way, oppression theorists are rightly concerned with idealizations that remove the presence of structural oppression, ignoring relevant differences between our forms of human life. Such a model will always be ideologically and ethically suspect. More importantly, such a model forecloses on the question of possibility. This is the crux of Vaihinger’s account. We model reality in order to better understand and manage it for ourselves. In this way, the question of what shared agency *really is at its core* misses the point. Shared agency allows us to manage our reality in better or worse ways. Investigating how shared agency emerges and finding ways to navigate those different forms should not foreclose on the question of what this second-order management could or should look like. Rather, such investigations should also start with the question: what do we want or need shared agency to be?³⁶

Conclusion

Finally, I conclude by arguing that social reality management is fundamentally about how people envision and enact the future. There is no shared vision of what the future should look like, nor is there, in an ontological sense, a single future. Practices of domination and oppression

³⁵ This is not to say that open-ended model pluralism necessarily takes developing large-scale models as its goal, but rather to highlight that there are cases in which certain forms of resistance will likely require such models, if only for reasons of efficacy. I take climate change to be such a case.

³⁶ This chapter does not ignore the question of who ‘we’ refers to. Also, this chapter pulls this question from literature on conceptual engineering, arguing that conceptual engineering constitutes a form of shared reality management.

are not oriented to the terms of the present; they are oriented to whose vision of the future is used in planning, who is and is not included in that future, and, if included, what the terms of that inclusion look like. Ontological forms of oppression also shift the terms through which people imagine their future, both individually and collectively, and also, of course, actively harm attempts of the oppressed to enact their visions of future, not only in terms of mere survival, but also flourishing. Without critical attention to how our partial and parochial models develop through past and present contours of our socio-ontological landscapes, we will fail to see how such models enact equally parochial pictures of the future. This is ultimately why idealizations of agent-neutrality, social ontology, and epistemology are ideologically and ethically concerning. Such idealizations, in virtue of their appeals to sameness, not only obscure the present realities of oppression, but ignore how shared agency, as a form of planning, as a form of social reality management, is fundamentally about what our futures look like and on whose terms.

Chapter I

Methodology and Idealizations in Shared Agency Research

Section I. The ‘Social’ in Social Ontology

Social ontology is the study of our social world, particularly in how people jointly create and maintain such worlds and the various phenomena within them. Shared agency, a subfield of social ontology, is an investigation of collective actions; actions that require, at minimum, the coordination of two or more people. Philosophical literature on shared agency, however, often develops models for how people successfully share agency by first removing many of the social, cultural, political, and hermeneutical differences that significantly shape who we are as agents and our capacity to act, both individually and collectively. By removing these features, theorists are able to offer basic models for shared agency that can be applied widely or, in some cases, universally. Such models do not claim to reflect social reality accurately nor in its full complexity, but rather operate as if such features did not exist in order to develop a model for shared agency that both (1) has a wide application, and (2) does not hinge on the contingent features of our social world and ourselves as agents. Operating similarly to the methodology of ideal theory in moral and political philosophy,¹ such models are applied in a top-down manner,

¹ John Rawls, *A Theory of Justice, revised edition*. (Cambridge, MA: Belknap Press of Harvard University Press, 1999). In *Theory of Justice*, Rawls begins from a methodological framework known as *ideal theory*. In political philosophy, ideal theory is the claim that in order to arrive at an understanding of what justice is we must begin from favorable assumptions about agents, their capacities, and their environment. Famously, this appears as the justificatory framework of his *original position*, wherein agents, who have no substantive identities, interests, or personalities and are removed from any social, political, historical, or geographical circumstances, will all agree to the two principles of justice. Ideal theory as a methodology is generally used by moral and political philosophers, serving as a justificatory framework for the set of normative principles in question. However, ideal theory can be understood more broadly, as it starts from a set of favorable assumptions about a phenomenon in order to better understand it. Such assumptions concern both the background conditions for the phenomenon, as well as the agents who participate in or are affected by it. The underlying commitment is that if we can understand how a particular phenomenon works under favorable, or ideal, conditions, then we are in better position to understand how it operates under unfavorable conditions. As a methodology, ideal theory functions much like conducting research in a lab environment. The researcher can control for (or eliminate) all factors that are regarded as irrelevant or secondary to

allowing for the removed features to be reintroduced through their non-ideal applications. In this way, much of the standard analytic literature on shared agency develops models that rely heavily on the methodological tool of idealization. Broadly construed, idealization is any form of abstraction, generalization, or simplification that introduces a false fact into a model and, in doing, so aids in the use of the model for the purposes of understanding, predicating, or otherwise managing the phenomenon under investigation. While idealization is not unique to ideal theory, many shared agency theorists employ idealizations in a similar manner to ideal theorists: by reducing agents to a set of fundamental or necessary features and thus removing the variety of differences between agents in order to arrive at a more universally applicable model.²

The use of idealizations within ideal theory has been a long-standing source of controversy, and heavily criticized in particular by non-Rawlsian,³ non-ideal theorists and oppression theorists more broadly. Following Lisa Tessman's broad definition, oppression theorists are: "all those, including most feminist and critical race theorists, who understand the phenomenon of oppression to be an important and unjust structural feature of worlds about which they do some sort of normative theorizing, such as ethical, social, and/or political

the phenomenon in question in order to arrive at the purest, or best, understanding of it. Since the phenomenon in question is shared intentions, then the general claim is that by removing the particularities of agents and the specific contexts in which agents act, we can arrive at the best understanding of how people successfully share intentions.

² Such theorists include but are not limited to: Christopher Kutz, *Complicity: Ethics and Law for a Collective Age* (Oxford: Oxford University Press, 2000); Michael Bratman, *Shared Agency: A Planning Theory of Acting Together* (Oxford: Oxford University Press, 2013); Raimo Tuomela, *Social Ontology: Collective Intentionality and Group Agents* (Oxford: Oxford University Press, 2013); Margaret Gilbert, *Joint Commitment: How We Make the Social World* (Oxford: Oxford University Press, 2014).

³ Since John Rawls' initial understanding of ideal theory included non-ideal theory as well, there are two kinds of non-ideal theorists; (1) those who regard themselves as taking up the non-ideal application of Rawls' ideal project; and (2) those who reject the methodology of ideal theory altogether. See, e.g., Lisa Tessman, ed., *Feminist Ethics and Social and Political Philosophy: Theorizing the Non-Ideal* (New York: Springer, 2009).

theorizing.”⁴ Many oppression theorists, most notably Charles Mills,⁵ argue that ideal theory (and its use of idealizations in particular) actively obscure mechanisms of oppression, offering political and moral theories that reflect the parochial values of those developing the theory, particularly the interests of dominate or non-oppressed groups within society.⁶ In response, such theorists have largely eschewed idealizations altogether, taking up concrete projects that begin from the very features that ideal theorists tend to remove: race, gender, sexuality, class, ability, and so forth.

In this chapter, I begin my methodological critique of standard analytic literature on shared agency. Drawing on Hans Vaihinger’s *The Philosophy of ‘As If,’*⁷ I argue that the ideal methodology employed relies on three standard idealizations that remove agents’ social features and the social world in which agents are embedded from their models for how two or more agents successfully shared agency: (1) the idealization of agent neutrality, (2) an idealized social ontology, and (3) an idealized epistemology. Drawing on non-ideal, oppression theory, I argue that these idealizations provide guidance only for a counter-factual world, not the actual world, obscuring the dynamic and contextual processes of forms of structural oppression and their impacts on agents.

⁴ Lisa Tessman, *Moral Failure: On the Impossible Demands of Morality* (Oxford: Oxford University Press, 2014), 176n2.

⁵ Charles W. Mills, “‘Ideal Theory’ as Ideology,” *Hypatia* 20, no. 3 (Summer 2005): 165-84.

⁶ See, also: Lisa Schwartzman, *Challenging Liberalism: Feminism as Political Critique* (University Park, PA: Pennsylvania State Press, 2006); Carol Pateman and Charles Mills, *Contract & Domination* (Malden, MA: Polity Press, 2007); José Medina, *The Epistemology of Resistance: Gender and Racial Oppression, Epistemic Injustice, and Resistant Imaginations* (Oxford: Oxford University Press, 2013).

⁷ Hans Vaihinger, *The Philosophy of “As If”: A System of the Theoretical, Practical, and Religious Fictions of Mankind*, trans. C.K. Ogden (New York: Harcourt, Brace, and Co., 1924.)

However, I also argue that non-ideal oppression theorists are wrong to reject idealization as a valuable methodological tool in the cases where idealizations provide enhanced understanding and management of our actual world. I show why oppression theorists are right to reject the use of idealizations in ideal theory and thus, by extension, the similar use in accounts of shared agency, but also wrong to reject idealizations altogether. In developing models of shared agency, idealizations are a necessary tool, but not all forms of idealizations are methodologically, politically, or ethically justified, including those that result in the development of universal or otherwise highly general models for shared agency. Vaihinger argues that we develop models both for the purposes of understanding a phenomenon *and* for managing it. If our purpose is to understand how particular forms of shared agency arise – be they successful or not – in order to enhance our capacity to more successfully share agency, especially in cases of shared resistance, then our models must develop through dialogical engagement with actual, concrete conditions. The highly general and top-down models found throughout the analytic literature on shared agency, in contrast, idealize away from complexity, excluding the presence and impact of, what Patricia Hill Collins refers to as, a *matrix of domination*⁸ in shaping *agential identity*,⁹ social groups, interpretative communities and the role of power in knowledge production,¹⁰ and the possibilities for sharing agency between those who are positioned differently within such systems of power. The idealizations used by standard models with the

⁸ Patricia Hill Collins, *Black Feminist Thought: Knowledge, Consciousness, and the Politics of Empowerment* (New York: Routledge, 1991), 225-230. *Matrix of domination* refers to the way in which systems of power necessarily interlock (i.e., exist in co-formation with one another) such that a single-axis approach to understanding or resisting forms of domination is insufficient.

⁹ Robin Dembroff and Catharine Saint-Croix, “‘Yep, I’m Gay’: Understanding Agential Identity,” *Ergo* 6 (2019): 571-599.

¹⁰ Patricia Hill Collins, *Intersectionality as Critical Social Theory* (Durham, NC: Duke University Press, 2019), 87-156.

analytic literature are also either implicitly or explicitly normative;¹¹ the former in virtue of the kinds of evaluative and ontological assumptions reflected within the idealizations used. Thus, my broader methodological concern is an ideological and ethical one. Over-idealized, highly general models offer a guide for how people ought to share agency without attending to the impact of those relations of power that shape who we are, individually and collectively. In doing so, such models attempt to *manage* how people share agency through overly general and uncritically parochial models.

If shared agency projects are one significant way in which we shape and/or manage our social reality, then highly general, over-idealized models for how people successfully share agency are functionally tyrannical, giving users of the model the ability to drastically shape: (1) what counts as a form of shared agency; (2) how we share agency together; (3) who is regarded as the right kind of agent for shared forms of agency, which I refer to as a form of *agency monism*; (4) the extent to which previous forms of social reality management (or forms of mismanagement)¹² are reiterated through our current and future shared projects – also called *ontic injustice*.¹³ When models for shared agency use idealizations that actively exclude the impact of structural oppression, they risk adopting the unjust outcomes of previous forms of social reality management (e.g., conceptual schemas around race, gender, etc.). This makes it impossible, among other things, to use these models for shared agency and joint action aimed at resisting structural oppression. That is, if we already exist in an unjust social ontology, then

¹¹ Some analytic models are explicitly normative, while others are implicitly normative insofar as their ontological and evaluative assumptions included in the model function normatively (as opposed to purely descriptively).

¹² By ‘mismanagement,’ I refer broadly to the fact that we already live within an unjust social ontology that shapes our agency, conceptual schemas, and levels of knowledge or ignorance. This unjust social ontology is partially the product of previous attempts to manage the reality of our social worlds in unjust ways.

¹³ Katharine Jenkins, “Ontic Injustice,” *Journal of the American Philosophical Association* (forthcoming).

social ontological resistance must address the forms of ontic injustice that already constitute features of our existence, our communal experiences, and interactions between members of similar or dissimilar communities. When standard models fail to critically assess whether their methodological approach, use of idealizations, or models themselves reflect or reinscribe forms of ontic injustice, then the use of such models as a forms of second-order social reality management is likely to reinforce the co-constitution and continuation of unjust systems of power. More pressingly, since my particular interest in undertaking this methodological project is about the question of resistant forms of shared agency – those forms of shared agency developed within and between communities aimed at resisting interconnected systems of power – then such models necessarily do not provide the resources to either understand such forms of resistance nor enhance our capacity to undertake them together.

Thus, in this chapter, I provide an overview of Vaihinger's fictionalist methodology, highlighting in particular his discussion about the relationship among models, idealizations, and purpose. Using Vaihinger, I argue that non-ideal oppression theorists are right to reject contextually unjustified idealizations that only provide guidance for counter-factual conditions, but wrong to reject idealizations altogether. Non-ideal oppression theorists should not only embrace the methodological value of contextually justified idealizations, but also already employ such idealizations in order to provide understanding or action-guidance under actual, non-ideal conditions. My goal in making this point is ultimately to show in the proceeding chapters that certain forms of idealization can be practically useful for understanding and enhancing various forms of shared agency, but also resistant forms of shared agency in particular.

Section II. Vaihinger: Thought, Purpose, Models

In *The Philosophy of “As If”: A System of the Theoretical, Practical and Religious Fictions of Mankind*,¹⁴ Hans Vaihinger shows how a fictionalist, or proto-pragmatist,¹⁵ methodology underlies science and philosophy. Inspired by Immanuel Kant’s theory of antinomies and prioritization of practical reason,¹⁶ Vaihinger argues that fictions, or the method of ‘as if,’ play a tremendous role in science, philosophy, and our everyday life: “Fictions are never verifiable, for they are hypotheses which are known to be false, but which are employed because of their utility.”¹⁷ Fictions, or idealizations, are theoretical falsehoods with practical value, particularly in developing a model, i.e., a representation of some process or phenomenon. When developing a model, the model has a particular purpose. For instance, one might internally model the same street in incompatible ways, depending on the purpose of the model: walking or driving on this street. Incompatible models can sometimes be rendered compatible by solving the underlying fragmentation, but this is not true in all cases; however, incompatible models still are valuable, depending on the purpose for which we need the model. For example, P.F. Strawson’s account of participant–reactive attitudes highlights how we switch between the objective stance, a deterministic picture of human behavior, and the participant stance that characterizes our

¹⁴ Vaihinger, *The Philosophy of “As If.”* Noted with full title in text.

¹⁵ Vaihinger, *The Philosophy of “As If,”* viii. “Fictionalism does not admit the principle of Pragmatism which runs: ‘An idea which is found to be useful in practice proves thereby that it is also true in theory, and the fruitful is thus always true.’ The principle of Fictionalism, on the other hand, or rather the outcome of Fictionalism, is as follows: ‘An idea whose theoretical untruth or incorrectness, and therewith its falsity, is admitted, is not for that reason practically valueless and useless; for such an idea, in spite of its theoretical nullity may have great practical importance.’ But though Fictionalism and Pragmatism are diametrically opposed in principle, in practice they may find much in common. Thus both acknowledge the value of metaphysical ideas, though for very different reasons and with very different consequences.”

¹⁶ Immanuel Kant, *Critical of Pure Reason*, trans. Marcus Weigelt (New York: Penguin Books, 2007).

¹⁷ Vaihinger, *The Philosophy of “As If,”* xlii.

ascriptions of moral responsibility in everyday life.¹⁸ These models are incompatible, but the potential truth of determinism cannot prevent our social practices of holding each other accountable, because it is a fundamental part of human social life. Because humans must thus navigate the world through fundamentally incompatible models combined with the fact that there are limitations to human cognition, we develop tools for the investigation of reality that reflect the kinds of limited creatures we are and the needs we have. Vaihinger argues that idealizations, or the method of ‘as if,’ is one of our most significant and widely used tools.

Breaking from Kant, Vaihinger argues that the limitation of thought and knowledge is not particular to humans but rather embedded in the nature of thought itself. Drawing particularly on Arthur Schopenhauer’s metaphysics of will and Wilhelm Wundt’s theory of the heterogeny of purpose,¹⁹ he argues:

...Such limitations are part of the nature of thought itself; that is to say, if there are higher forms of mind, these limitations will affect them and even the highest Mind of all. For thought originally serves the Will to Life as a means to an end, and in this direction also fulfils its function. But when thought has broken loose from its original aim, according to the Law of the Preponderance of the Means over the End, and has become an end in itself, it sets itself problem to which it is not equal because it has not developed for this purpose; and finally the emancipated thought sets itself problems which in themselves are senseless...

In a basic way, Vaihinger relates the notion of purpose in a model to that of purposeful activity in organisms – in this case of humans, this activity is thought:

All purposeful activity manifests itself in seeking out, collecting, or producing the necessary and serviceable means for the attainment of its object. The organic activity of thought also manifests its purposeful nature in exerting itself to attain its aims by all the means at its command.²⁰

¹⁸ P.F. Strawson, “Freedom and Resentment,” in *Freedom and Resentment and Other Essays* (New York: Routledge, 2008), 1- 28. For instance, Strawson’s account of the objective stance and reactive attitudes provides a way to manage two potentially incompatible models, namely determinism and moral responsibility.

¹⁹ Arthur Schopenhauer, *The World as Will and Representation* (New York: Dover Publications 1969); Wilhelm Wundt, *Ethics, An Investigation of the Facts and Laws of Moral Life* (London: Sonnenschein & Company, 1897).

²⁰ Vaihinger, *The Philosophy of “As If,”* 6.

The connection between purposeful behavior generally and human thought specifically, for Vaihinger, is that the structures of thought develop in accordance with our need to better navigate – and thus manage – our external and internal worlds:

It must be remembered that the object of the world of ideas as a whole is not the portrayal of reality—this would be an utterly impossible task—but rather to provide us with an *instrument for finding our way about more easily in this world*.²¹

Vaihinger notes that this task is already marked by contradiction from the beginning.

Distinguishing between what Maurice Merleau-Ponty would later call the *world of science* and *the world of perception*,²² Vaihinger argues that the *world of motion* and the *world of consciousness* are contradictory concepts of reality. On the one hand, his account of thought reduces the world to one of sensations (in which humans must model reality in light of, what are often contradictory, sensations, which means we are modeling reality in relation to the kinds of organisms we are and needs we have);²³ on the other hand, he relies on a concept of reality from the natural sciences in which reality is reduced to small entities and their movement.²⁴

Responding to potential critics of this fundamental incompatibility within his theory, Vaihinger argues that this is precisely the point of his methodology:

²¹ Vaihinger, *The Philosophy of “As If,”* 15. Italics in original text.

²² Vaihinger, *The Philosophy of “As If,”* xliii-xliv. Vaihinger refers to this at the *world of motion* (particularly in relation to the physical sciences) and the *world of consciousness* (xliv). I find Merleau-Ponty’s language to be more useful in highlighting this distinction, but Merleau-Ponty and Vaihinger have slightly different ways of conceiving of this distinction (and Merleau-Ponty rejects Vaihinger’s slightly behaviorist account of sensation). However, the similarity here between these two theorists is noteworthy and would be a fruitful point of further investigation into the relationship between these two contradictory concepts of reality. See: Maurice Merleau-Ponty, *The World of Perception*, trans. Oliver Davis (Routledge: New York, 2008).

²³ One might liken this to J.J. Gibson’s discussion of the relationship between ecological niches and affordances (i.e., possibilities for action). See: J.J. Gibson, “The Theory of Affordances,” in *Perceiving, Action, and Knowing: Towards an Ecological Psychology*, eds. John D. Bransford and Robert Shaw (Hoboken, NJ: John Wiley & Sons, 1977), 127-142.

²⁴ Vaihinger, *The Philosophy of “As If,”* xliii-xliv.

Philosophical analysis leads eventually, from an epistemological standpoint, to sensational contents, and from a psychological to sensations, feelings and strivings or actions. Scientific analysis leads to another concept of reality, to matter, and the smallest constituents and motions of matter. Naturally it is impossible for the mind as such to bring these two spheres of reality into a rational relation, although in institution and experience they form a harmonious unity.²⁵

In this way, Vaihinger argues that our models of reality must start from the kinds of organisms we are; ones who are not merely capable of holding contradictory ideas but who must necessarily navigate the world in terms of that capacity. The philosophical ideal to explain everything rationally thus fails to engage with the fundamental irrationality that shapes thought, in which there will never be a rational relation between these two concepts of reality, especially given the cognition limitations of the human organism.²⁶

In *As If: Idealization and Ideals*,²⁷ Kwame Anthony Appiah offers an extended analysis of Vaihinger's methodology, noting that if we were *Cognitive Angels*—those rational beings incapable of any logical contradiction in their concepts of reality—then we would not even be in the position to experience this contradiction. That is, as Appiah writes,

It is our imperfection that allows us to work, not with a single picture of the world, but with many. And because they are incompatible with one another — because they cannot all be true—we have to be able to keep them separate if we are not to be drawn into incoherence. I *can* think of the earth as spherical and as ellipsoidal, for different purposes; what I *cannot* reasonably do is think of it both ways for the same purpose and at the same time. *Our knowledge of reality is held, then, in pictures of the world, each of which has something wrong with it but is good enough for some purposes.*²⁸

²⁵ Vaihinger, *The Philosophy of "As If,"* xlv.

²⁶ Vaihinger, *The Philosophy of "As If,"* xxix. As Vaihinger writes, "I therefore consider it to be a lack of sincerity in most systems of philosophy, that they tried more or less to hide the irrational side."

²⁷ Kwame Anthony Appiah, *As If: Idealization and Ideals* (Cambridge, MA: Harvard University Press, 2017).

²⁸ Appiah, *As If*, 110-111. Italics mine.

Thus, a rationally unified, singular picture of reality is not possible. However, even if we develop two contradictory models of a phenomenon for different purposes, the joint use of such models can sometimes be useful in its own right, even if they have no rational relation, which is why Vaihinger argues that the unity of contradictory models is only in experience – that is, the joint use of models is, in fact, useful for us, even if rationally ought not to be.

Thus, in creating models of various features of reality, our models will often be contradictory but these contradictions generally track differences in purpose. Our purposes, in turn, reflect our practical and theoretical needs and values. The way we model features of reality is already a function of what features people take to be most pressing in terms of understanding and management. For instance, relatively little is known about menstruation, which is unsurprising given long-standing androcentric bias within biological and medical research, and this impacts our ability to manage this biological process (e.g., pain reduction).²⁹ Moreover, how we model features of reality is not merely reflective of particular values or interests but also reflective of the kinds of conceptual schemas in which the modelers exist, impacting both hypothesis development and what one accepts as evidence. That is, a particular gender schema impacts how one researches the phenomenon of gender, the kind of model one develops, the purpose of the model, and the scope in which the model is applicable. Moreover, one's conceptual schema(s) impacts how one perceives the phenomenon under investigation in better or worse ways, which not only impacts how one develops a model but also the extent to which the model is deemed useful in fulfilling its purpose. This is why interventions by feminist philosophers of science and feminist social epistemologists are so critical: humans do not all have the same needs and values (with exception to certain basic needs attached to the value of

²⁹ Caroline Criado Perez, *Invisible Women: Data Bias in a World Designed for Men* (New York: Abrams Press, 2019), 174.

survival), and our epistemic situatedness means that we often fail to see this fact, which is, in turn, reflected in how we construct models of reality in order to understand and manage it.

Since a single model of reality is both insufficient and impossible, a single model of phenomenon, i.e., model monism, is insufficient as well. Appiah's analysis of Vaihinger's theory of idealizations further extends his account from the realm of science, ethics, and religion to the social sciences and political philosophy more broadly. When we model human behavior, model monism is insufficient because: (1) we often need to model the same behavior for different purposes; (2) no single model will be able to adequately account for all the relevant forms of human difference that often shape our behavior in contextually dependent ways; (3) the same behavior can reflect different forms of purpose or meaning depending on the context, and thus models cannot rely on context-independent idealizations; and (4) modeling and categorizing human behavior can change the behavior, which Ian Hacking refers to as "the looping effect."³⁰

Model monism (or *constrained model pluralism*)³¹ in this case would require idealizations so general and abstract that the model would be useless for increased understanding. Moreover, since models are inherently value-laden, model monism (or constrained model pluralism) imposes a parochial set of values under the guise of value-neutrality and impartiality, and thus attempts to manage human behavior while, at the same time, evading critique of the values used to do so. Modeling forms of shared agency is no different. How people share agency, our reasons for doing so, how 'sharing' and 'agency' are understood, the terms of success or failure, who is a potential agent with whom another can share agency are

³⁰ Ian Hacking, *The Social Construction of What?* (Cambridge, MA: Harvard University Press, 2012). It is worth noting that Hacking requires that people respond to this process of categorization and the creation of new human kinds, otherwise the looping effect does not occur.

³¹ See Chapter V for full analysis.

all contextually dependent (e.g., on: culture, agential and group identities, geographical location, conceptual schemas, and so forth). Model monism (or constrained model pluralism) cannot provide an increased understanding of how people share agency in these context-dependent ways, but, as Hacking notes, the application of such a model can affect our actual social relations and conceptual schemas (while, at the same time, uncritically adopting the outcomes of previous and unjustified attempts at social reality management). Thus, as I will argue in Chapter V, *open-ended model pluralism* offers a methodologically superior approach for understanding the myriad ways shared agency arises in the world while ensuring critical attention to kinds of values, metaphysical assumptions, and conceptual schemas taken up by the individual models, insofar as these models are also used to manage shared agency.

Section III. Idealizations, Ideal Theory, and Oppression

Vaihinger's methodological fictionalism provides a way of thinking about the role of idealizations in how we develop models of reality. A model is a representation of some process or phenomenon that relies on some form of idealization in order provide a better understanding of that process or phenomenon *as well as* allow the model's users to manage that phenomenon in accordance with a particular set of values or needs. Models can be more or less idealized; and less idealized models require significantly more data points. However, even models that rely on a high number of data points (or, better put, tend to include more complexity) still rely on idealizations. For instance, both Vaihinger and Appiah point to Newton's laws of motion, which provide a model for understanding how force, mass, momentum, and acceleration work, but does so using idealizations such as a frictionless plane.³² In this way, our models of reality and

³² Appiah, *As If*, 26, 50, 77.

particular phenomena therein rely on contextually justified idealizations. Idealizations are contextually justified when the model in which they are used: (1) provides an enhanced understanding of the phenomenon under investigation, and (2) that this understanding enables its users to better navigate, or manage, the phenomenon. Whether a model works depends on what the model sets out to accomplish, and, in some cases, the purpose of a particular model is open to interpretation.

In contemporary moral, social, and political philosophy, the use of idealizations is most commonly associated with ideal theory. Ideal theory, broadly defined, is a methodological framework for arriving at a normative principle (or set thereof) which begins from a set of favorable idealizations about agents, their capacities, and their environment. By deriving normative principles from ideal circumstances, the idea is that such principles can then be generally applied universally in a top-down manner to non-ideal contexts. As the story goes, this ensures that our normative principles do not hinge on the contingent features of any particular social ontology and thus are appropriately universal. In this vein, there are two kinds of non-ideal theorists. First, Rawlsian non-ideal theorists focus on how to apply ideal theory's principles to non-ideal contexts. Second, non-Rawlsian, non-ideal theorists reject ideal theory for methodological and/or ideological reasons.

Many of the standard models for shared agency tend to function similarly to ideal theory's methodology for the development of normative principles. First, one develops a normative model by excluding the contingent features of any particular society such that the model can apply universally (or at least in a highly generalized way) to all societies. Models for shared agency operate similarly by providing a model for how two or more people shared agency with a narrow set of success conditions, and such models are taken to apply universally or, at

least, in a highly general manner. Such models thus provide a single or small set of models for capturing how people, regardless of context, successfully share agency.

Non-ideal conditions are added back in during the later stages when the model is directly applied to a particular context. In this way, deviations from the model are not taken as evidence that the model is insufficient, but rather such deviations are signs of failure to approximate the ideal for shared agency imposed by the model. This is because such models for shared agency are not merely descriptive but also normative, setting an ideal both for individual human agency and the conditions by which such agents can successfully share intentions. This means that models for shared agency rely on idealizations that offer a highly simplified, oppression-free social ontology and remove any features from agents that might be regarded as contingent or overly particular (thereby assuming that human agency can be defined by a set of universally shared features in which the prioritization of those features is also shared). Moreover, such accounts idealize human cognitive capacities through idealizations that: (1) function similarly to Appiah's notion of Cognitive Angel; (2) ignore how differences in conceptual schemas significantly impact our cognition and hermeneutical frameworks, which shape how we understand our social worlds and others within them; and (3) assume an asocial epistemological framework in which our knowledge of others' intentions and understanding of the creation and implementation of shared goals is not fundamentally impacted by problems like epistemic injustice, forms of pernicious listening and unjustified failure to give speakers' claims proper uptake, and hermeneutical impasses.³³

³³ See Miranda Fricker, *Epistemic Injustice: Power and the Ethics of Knowing* (Oxford: Oxford University Press, 2007); Kristie Dotson, "Conceptualizing Epistemic Oppression," *Social Epistemology* 28, no. 2 (2014): 115-138; Luvell Anderson, "Hermeneutical Impasses," *Philosophical Topics* 45, no. 2 (2017): 1-19.

In these ways, standard accounts of shared agency rely on an ideal theory-like methodology when they ideal *away* from forms of structural oppression, processes of social construction and differences in conceptual schemas, and ignore differences between agents. As a methodology, ideal theory aims to provide a highly generally or universally applicable principle or model, and many standard accounts of shared agency also retain this goal. Moreover, shared agency accounts that rely on this sort of methodology tend to define shared agency in terms of a narrow set of success conditions, meaning that anything that does not match these criteria are either failed attempts at shared agency or simply do not qualify as shared agency at all (i.e., not even as a failed attempt). The models for shared agency produced by such accounts thus tend to rely on three forms of idealization: (1) the idealization of agent-neutrality, which oversimplifies human difference; (2) an idealized epistemology (and cognitive sphere, i.e., Cognitive Angels); and (3) an idealized social ontology that excludes the presence of oppression and similarly treats difference in social ontologies as superfluous rather than constitutive. This third idealization functions as a form of decontextualization; if models are not context-dependent, they can apply to a wider range of contexts. Given the connections between ideal theory in political and moral philosophy and its similar occurrence throughout literature on shared agency, the critique of ideal theory and its use of idealizations by oppression theorists provides valuable insight as to why the models that result from such idealizations are both inferior in terms of use-value and ethically concerning.

Oppression theorists tend to be non-ideal theorists of the non-Rawlsian sort and reject ideal theory for three general reasons. First, ideal theory's use of idealizations only provides normative guidance for the counter-factual worlds they presuppose and thus operate as unattainable ideals for agents' capacities. Onora O'Neill distinguishes between abstractions and

idealizations, arguing that idealizations necessary fail as a method.³⁴ O'Neill argues that abstraction is a necessary feature of theorizing because theory involves some level of generalization. Abstraction is a function of omission that leaves out certain predicates in the descriptive or theoretical content. In this way, abstraction is not unique to ideal theory but rather a central feature of practical reasoning. In contrast, idealizations offer new predicates, often relying on counterfactual stipulations. Idealizations involve idealized, hypothetical agents, whose capacities are unlike actual human agents. O'Neill objects to the use of idealization in moral and political theory because they build a substantial amount of false information about human agents. This false picture of agency is concerning because no actual agent can achieve it, creating a problem of justification. Any theory that relies on idealizations attempts to justify their use by pointing to a set of metaphysical assumptions but often do not (or cannot) supply the right justification for those assumptions. Unlike omissions, idealizations undermine a theory's application potential and prevent the transition from ideal to non-ideal theory. Moreover, these idealization function normatively. While humans will never be the Cognitive Angels that such idealizations presuppose, these idealizations still function as a guiding ideal for human agency and capacities. Building off O'Neill's critique, Lisa Schwartzman goes a step further, arguing that the method of abstraction itself assumes that we sufficiently understand the phenomenon in question in order to know what to omit when abstracting.³⁵ For Schwartzman, abstracting away from the presence of structural oppression presupposes that we sufficiently understand the

³⁴ Onora O'Neill, "Abstraction, Idealization, and Ideology in Ethics," in *Moral Philosophy and Contemporary Problems*, ed. J.D.G. Evans (Cambridge, UK: Cambridge University Press, 1987), 55-69; Onora O'Neill, "Ethical Reasoning and Ideological Pluralism," *Ethics* 98, no. 4 (July 1988): 705-722.

³⁵ Schwartzman, *Challenging Liberalism*, 5-9, 160, 171.

impact of it on the development and expression of human agency, both individually and jointly, when we, in fact, do not.

However, Vaihinger argues that idealizations include methods like abstraction, because an omission itself implicitly introduces a new predicate. In this way, Schwartzman's extension to abstraction better captures the scope of idealizations than O'Neill. However, both theorists capture a problem with how ideal shared agency theorists use idealizations. For Vaihinger, idealizations are useful insofar as they provide us both with a better understanding of the phenomenon under investigation *and* allow us to better manage the phenomenon. Idealizations that include Cognitive Angels cannot do this, because they introduce a predicate that can only tell us how agents in that counterfactual world can or should act. That is, the idealization of Cognitive Angels presupposes we are the kinds of beings who cannot hold two incompatible pictures of reality in our head at once when, in fact, that is exactly who we are. When a model uses this idealization, it only provides guidance for agents who are unlike ourselves and thus the model is not useful on Vaihinger's account. This does not mean we cannot use idealizations about human capacities or cognition, but rather such idealizations must start from the kinds of agents we are, which includes attending both to our own limitations as well as the social differences that impact the development and expression of our capacities. In this way, Schwartzman is right to be concerned with forms of abstraction that exclude relations of power and oppression, because they result in models that cannot provide us with an enhanced understanding of how power relations impact agents and social institutions, nor can it provide a way to manage (or effectively respond to) problems related to structural oppression.

Second, following O'Neill, Charles Mills argues that idealizations exclude the role of structural forms of oppression, particularly in shaping parts of our agential identities and our

social-structural positioning within these unjustified hierarchies, and thus tend to reflect values and purposes of those who largely benefit from oppression (even if unintentionally).³⁶ Mills distinguishes between two kinds of models: *ideal-as-idealized-model* and *ideal-as-descriptive-model*.³⁷ Newtown's laws of motion function as the latter, stipulating counterfactuals that enhance our understanding of the phenomenon in question under non-ideal, or actual, conditions. Ideal-as-idealized-model, however, only provides us with an understanding of how the phenomenon would work under those counter-factual conditions. Ideal theory uses idealizations as ideals in this way. As Mills writes, ideal theory "...either tacitly represents the actual as a simple deviation from the ideal, not worth theorizing in its own right, or claims that starting from the ideal is at least the best way of realizing it."³⁸ Ideal theorists often include: (1) an idealized social ontology that presupposes a metaphysically liberal conception of human agency, which excludes the presence of structural forms of oppression and the role of social groups in our social world, and often also over-idealizes the role of social institutions; (2) offers an impossible-to-obtain picture of human capacities and ignores how oppression further impacts the development and expression of their non-ideal counterparts; and (3) offers an idealized picture of human cognition in which "a general social transparency will be presumed," and which pays little to no attention to "the distinctive role of hegemonic ideologies and group-specific experiences in distorting our perceptions and conceptions of the social order."³⁹ Ideal theory's use of *ideal-as-idealized-models* thus cannot provide enhanced understanding about actual forms of injustice in our world and, by extension, cannot provide us with the tools for amelioration. When shared

³⁶ Mills, "'Ideal Theory' as Ideology."

³⁷ Mills, "'Ideal Theory' as Ideology," 166-167.

³⁸ Mills, "'Ideal Theory' as Ideology," 168.

³⁹ Mills, "'Ideal Theory' as Ideology," 169.

agency theorists employ a similar methodology with such idealizations, this means that such models for shared agency cannot attend to the impact of forms of structural oppression on agents, their membership in various social groups, and relations between agents with different social-structural positionings.

Third, and finally, one significant criticism of ideal theory more broadly is, as Mills notes, its tendency to use idealizations that presuppose a metaphysical liberal conception of human agency. This constitutes a form of, what I call, *agency monism*: the belief that human agency is, at its core, fundamentally the same and that the liberal picture of agency is the correct one. Ideal theory cannot function without agency monism, because ideal theory presupposes justice monism, a term introduced by Amartya Sen.⁴⁰ Justice monism is the claim that there is a single and universal picture of justice. For example, Rawls' idealization of the original position provides us with two basic principles of justice, developed by excluding contextual and non-ideal features of our social worlds, that can be applied to all societies.⁴¹ Contextual and non-ideal differences are added in secondarily during the application of the principles, instead of including such differences in the development of the principles themselves, and thereby limiting the application of the principles to the particular context in which they were derived (a method used more commonly within critical theory). Thus, the goal of ideal theory in its non-ideal applications is to bring all societies closer toward this monistic ideal.

The problem with agency monism is that it: (1) offers a contestable picture of agency that is not universally shared, which appeals to parochial values when defining the necessary or important features of agency (i.e., individual autonomy over communal autonomy); (2) ignores

⁴⁰ Amartya Sen, *The Idea of Justice* (Cambridge, MA: Harvard University Press, 2009).

⁴¹ Rawls, *A Theory of Justice*, 15-19.

how parochial conceptions of agency and agents themselves are significantly shaped by processes of social construction by assuming that any perceived variance is superficial or erroneous; and (3) fails to consider that parochial conceptions of agency fundamentally shape how agents experience and express their agency. When models for shared agency employ agency monism, they provide models for how shared agency ought to function rather than capturing the ways in which shared agency actually functions, unable to account for the differences in how agency develops particularly in metaphysically pluralistic ways.

Thus, oppression theorists askew ideal theory and idealizations under the assumption that the methodology use of idealizations is incompatible with projects aimed at concrete understanding and practical amelioration of structural forms of oppression. However, Vaihinger's fictionalism offers support for oppression theorists' criticisms concerning the role of idealizations in ideal theory, illuminates why oppression theorists themselves do not avoid idealization, and provides explanation as to why the idealizations of oppression theory differ from those of ideal theory. Broadly stated, our models of reality and particular phenomenon within rely on contextually justified idealizations. Idealizations are contextually justified when the model provides an enhanced understanding of the particular phenomenon under investigation and that this understanding enables its users to better navigate, or manage, the phenomenon in question.⁴²

First, contra O'Neill, Vaihinger argues that idealizations are useful fictions, or untruths, that do, in fact, provide us with useful models for understanding and managing features of reality. That is, a model of reality does not necessarily fail to be applicable if it includes new predicates that are false. Second, Vaihinger offers a different, more expansive definition of

⁴² That is, whether a model works is dependent on what the model is intended to accomplish.

idealizations than O'Neill. For Vaihinger, idealizations are any sort of as-if theorizing; that is, any time a theorist knowingly uses an untruth in order to arrive at a better understanding of an actual phenomenon and to better manage that phenomenon in relation to human life. In this way, idealizations include any form of abstraction, generalization, heuristics, approximations, reliance on simple cases, and hypotheses. On this model, all oppression theorists rely on idealizations of one form or another.⁴³ This is why Vaihinger argues that idealizations are a necessary part of theorizing. Humans are too cognitively limited to develop perfectly accurate models of reality because they would be computationally intractable. Idealizations are thus a tool for the non-ideal reality of human capacity. They are the way by which we develop better models of the world around us (and ourselves), providing us with an increased capacity to navigate, or manage, that world.

Third, and most importantly, Vaihinger's analysis of idealizations actually provides a valuable way to critique the relationship between idealizations and ideals. First, similar to O'Neill, Vaihinger argues that idealizations that can only provide us with a picture of how the world would work if the idealizations were true (that is, if the new predicate were true) are bad idealizations. They treat a fiction as real *and* as an ideal, providing models of the world that are, in fact, useless for enhanced understanding and management, but also harmful in imposing impossible ideals. On Vaihinger's analysis, idealizations need to be contextually justified and tailored to the purpose of the model in which they are being used. Moreover, the scope of a model also must be contextually justified. For example, a model for what a successful marriage looks like (if there could be such a thing) could not be more general than the form of marriage it is intended to explain and manage. That is, there is no decontextualized model for what a

⁴³ Tessman, *Moral Failure*, 176. Tessman notes that, for example, non-ideal oppression theorists often rely on idealizations such as the stipulation that a moral agent can act at all.

successful marriage looks like, because there is no generalized model for what a marriage is or looks like. It depends on the cultural, social, and political milieu in which the practice arises. Moreover, any such model would already presume that marriage is a valuable kind of relationship, and it is one that can be more or less successful. That is, any such model already implies a certain set of assumptions and value commitments. Examining a model's idealizations also requires us to examine the kind of values or assumptions implicit in the creation of such a model in the first place.

This leads to the second way in which Vaihinger's account is valuable: it offers a different approach to the relationship between idealizations and ideals. Idealizations need not be normative, but they sometimes are. That is, idealizations can valuably function as ideals in some cases. Following Vaihinger, Appiah argues that ideals can function as a useful untruth that guides our behavior, but that theorists still must justify the use of these ideals for their particular purpose. Moreover, since theorists can over-idealize or under-idealize, the goal is to attune the idealizations to the particular purpose, ensuring that they are not too general or too narrow. This is why oppression theorists have good reason to be critical of ideal theory – it over-idealizes and uses idealizations that are not valuable for the purpose of making a society more just.⁴⁴ For example, anti-oppression activists operate off ideals, but those ideals are often derived by first examining the non-ideal world, applied with a limited scope, and are rarely static (i.e., ideals often shift in tune with the successes or failures of a movement). In this way, Vaihinger's account of idealizations allows theorists to idealize from structural oppression rather than away

⁴⁴ Moreover, it is similarly not surprising that oppression theorists often gravitate to justice-enhancing theories, like Amartya Sen's, which develop idealizations from the non-ideal world in order to effectively address actual forms of injustice as they arise within and between different contexts. See Chapter III, V for related analysis.

from it. The purpose of an ideal is to be able to better manage ourselves and our world, but this does not require ideals to function universally nor statically.⁴⁵

Section IV. Concluding Remarks

My goal in this chapter has been to begin my methodological criticism of standard accounts of shared agency within the analytic literature by likening their methodology to that of ideal theory. In relying on an ideal methodology, I claim that the models arising from standard accounts employ three idealizations – an idealized social ontology, an idealized epistemology, and the idealization of agent neutrality. Such idealizations exclude the very social features of the world necessary to explain the varieties and possibilities for shared agency. However, non-ideal, oppression theorists have attempted to respond to the problems of an ideal methodology by eschewing idealizations altogether. Drawing on Hans Vaihinger’s fictionalist methodology, I have argued that the problem is not idealizations in and of themselves. Rather, the problem is idealizations that start by excluding complexity and in doing so only offer guidance for the counter-factual world they presuppose. Contextually-justified idealizations, as I will discuss further in Chapter III-V, can valuably enhance our understanding of our actual world by using idealizations as tool to manage complexity, rather than remove it, in order to more productively attend to the impact of structural oppression on our capacity to share agency, especially across different social-structural positionings within hierarchies of oppression.

⁴⁵ For further discussion about the role of ideals and idealizations, see: Chapters III-V.

Chapter II

Shared Intentions and Idealizations

Section I. The Question of Methodology

In order to investigate a phenomenon such as shared agency, there are questions about what constitutes an appropriate methodology and set of methodological tools. For this dissertation, instead of asking, “By what conditions do we define a successfully shared intention?,” my focus is on a more foundational question: “What are the appropriate methodological tools for investigation, and are those tools being used correctly?” Given that my goal is to understand how structural oppression impacts the possibility of shared agency, I focus specifically on the tool of idealization. Within the last century, the use of idealizations in subfields like moral and political philosophy has been a major source of debate, particularly surrounding the use of ideal theory. Idealizations, however, are not particular to these subfields, or even to philosophical methodologies, most notably underpinning forms of scientific investigation. However, the use of idealization has been increasingly controversial within social, moral, and political philosophy, especially for oppression theorists (in which I include myself) who frequently argue that idealizations necessarily obscure (or outright exclude) the presence and mechanisms of structural oppression. Thus, there are those who reject the use of idealizations within areas of social, moral, and political philosophy on the basis that they not only result in bad and inadequate theories, but are also ideologically loaded in ways that reflect the interests of those who benefit from structures of oppression. On the other hand, there are those that argue that idealizations are not merely impossible to avoid – that is, idealizations are a

necessary feature of theory and the way by which we learn about and navigate the world – but are also a deeply valuable tool for investigation, if used correctly.¹

I begin my methodological analysis from the latter position while incorporating concerns from the former: idealizations are unavoidable, even for oppression theorists.² However, my argument is not that we ought to reject idealizations altogether (as many oppression theorists argue) in the case of shared intentions. Rather, the focus of my dissertation is on the way in which prominent accounts of shared intentions employ idealizations and whether these idealizations are good or bad. The quality of an idealization is a matter of justification, namely whether theorists are justified in making particular idealizations within particular investigative contexts. In developing a model for shared intention and joint action, we rely on idealization for two goals: (1) to understand how shared agency works; and (2) to manage shared agency through by way of the model. If shared intentions and joint actions (i.e., shared agency more broadly) are the way through which we manage our social realities, then developing a model for shared intentions is about managing how we manage, or create, social realities.

¹ For the most recent and broadest example (and the focus of this chapter), see: Kwame Anthony Appiah, *As If: Idealization and Ideals* (Cambridge, MA: Harvard University Press, 2017). In the case of ideal theory, see, e.g.: Ingrid Robeyns, “Ideal Theory in Theory and Practice,” *Social Theory and Practice* 34, no. 3 (July 2008): 341-362; Zofia Stemplowska, “What’s Ideal about Ideal Theory?,” *Social Theory and Practice* 34, no. 3 (2008): 319-40; Laura Valentini, “On the Apparent Paradox of Ideal Theory,” *The Journal of Political Philosophy* 177, no. 3 (2009): 332-355.

² Lisa Tessman, *On the Impossible Demands of Morality* (Oxford: Oxford University Press, 2014), 175-203. Tessman argues that non-ideal theorists still engage in idealizations. In this case, the solution is not to avoid idealizations altogether. Rather, it highlights the limitations of a single form of normative theorizing, in Tessman’s case, and the need for simultaneously but different theoretical approaches to phenomenon in question. Tessman’s argument here rests on her claim that normative theorist ought to ask themselves, “What do we want from a normative theory?” (i.e., what do we want a theory to do?). As I will discuss in Chapter III, this connects to Appiah’s argument that different and contradictory models of the same phenomenon can offer enhanced understanding of the phenomenon and provide different ways of managing the phenomenon and our relationship to it (Appiah, *As If*, 1-54).

In this second chapter, I show that three prominent accounts of shared intentions – those offered by Michael Bratman,³ Margaret Gilbert,⁴ and Christopher Kutz⁵ – do employ idealizations and detail what those idealizations are. This chapter sets up my further claim in Chapter III that these theorists rely on unjustified, or bad, idealizations. Again, the problem of unjustified idealizations is not merely that they result in less useful models but also that certain idealizations can reflect parochial values and ontological assumptions (particularly those of structurally dominate groups or who have a relatively high access to power within systems of domination), and this raises ideological concerns.

The layout of this chapter is as follows. In Section II, I provide a literature review for three of the most prominent accounts of shared agency – those of Michael Bratman, Margaret Gilbert, and Christopher Kutz. In Section III, I turn to the question of methodology and analyze how each theorist employs the methodological tool of idealization. Drawing on Kwame Anthony Appiah’s recent work, *As If: Idealization and Ideals*,⁶ I explain broadly what an idealization is, the relationship between idealizations and abstractions, and why idealizations are useful (and unavoidable). Appiah’s work highlights the methodological value of idealizations by also arguing against their misuse, which can be referred to simply as bad idealizations. My analysis here primarily aims to highlight the particular use of this methodological tool in the works of Bratman, Gilbert, and Kutz.

³ Michael E. Bratman, *Shared Agency: A Planning Theory of Acting Together* (Oxford: Oxford University Press, 2013).

⁴ Margaret Gilbert, *Joint Commitment: How We Make the Social World* (Oxford: Oxford University Press, 2014).

⁵ Christopher Kutz, *Complicity: Ethics and Law for a Collective Age* (Cambridge, UK: Cambridge University Press, 2010).

⁶ Appiah, *As If*.

Section II. Shared Intentions

The subfield of collective intentionality investigates how intentional states can be shared between two or more people, as well as how a group can be an intentional agent. At the level of individual human agents, our intentions inform and regulate our individual actions. Collective intentionality is thus the study of how intentions structure and produce collective actions.

Collective actions can be broken two into two types: joint actions and group actions. Literature on group intentionality focuses on whether a group can be an intentional agent in its own right, how a group develops intentional states, and to what extent a group can be moral agent.⁷

Examples of group intentionality include the actions of highly organized groups, such as a corporation or university. Shared intentionality, on the other hand, focuses on the question of what kinds of intentional states inform joint actions (i.e., an action undertaken by two or more agents). Examples of shared intentionality focus on joint actions, such as dancing the tango or performing in a jazz quartet. The relationship between intentions and joint action, in this case, becomes a question of whether coordinated, individual intentions can sufficiently explain how joint actions work, or if joint actions require a different kind of intentional state – one that is more strongly shared in some way.

My focus in this dissertation is on shared intentionality and joint action. I examine three prominent models – those of Michael Bratman, Margaret Gilbert, and Christopher Kutz – in order to analyze the idealizations used in each and to what extent such idealizations are viable. My reason for choosing these three models is that they sufficiently highlight the range of theories that currently exist within the subfield, providing a basis for broader application of my analysis.

⁷ See, e.g.: Christian List and Philip Pettit, *Group Agency: The Possibility, Design, and Status of Corporate Agents* (Oxford: Oxford University Press, 2011); Raimo Tuomela, *Social Ontology: Collective Intentionality and Group Agents* (Oxford: Oxford University Press, 2013); Deborah Tollefsen, *Groups as Agents* (Cambridge, UK: Polity Press, 2015).

By capturing this range, my hope is that my analysis of their particular idealizations can be extended to other models not covered here. In this way, my focus is not intended to be exhaustive but rather to show the array of idealizations that shared intention theorists employ. Models of shared intention do not all employ the same idealizations, but my goal is to show that there are deep similarities between the kinds of idealizations used in these prominent accounts, which has some bearing on how many other models of both joint and group action use idealizations.⁸

i. Michael Bratman: Shared Planning Agency

For Bratman, our capacity to plan and act together takes the form of: I intend that we *J*.⁹ This account reflects his broader commitment to an augmented form of methodological individualism, in which his model for shared agency builds directly upon his model for individual intention and planning.¹⁰ In order to understand Bratman's account of shared agency, it is first necessary to discuss his account of individual agency. For individual agency, Bratman's *planning theory* of intention regards planning as a central feature of our temporally extended, rational agency, which allows us to develop and realized complex goals through future-oriented intentions.¹¹ As Bratman writes:

We form future-directed intentions as parts of larger plans, plans which play characteristic roles in coordination and ongoing practical reasoning; plans which

⁸ To be clear, I believe my analysis can be extended to other accounts, but I do not have the space here to develop those arguments.

⁹ Michael E. Bratman, "Shared Intention," *Ethics* 104, no. 1 (1993): 97-113; Michael E. Bratman, *Shared Agency: A Planning Theory of Acting Together* (Oxford: Oxford University Press, 2013).

¹⁰ Bratman, *Shared Agency*, 11-15. This individualism is augmented insofar as Bratman's builds up from his account of individual planning agency by "highlighting special contents of and interrelations between the plan states of such individual agents" (11).

¹¹ Michael Bratman, *Intentions, Plans, and Practical Reason* (The David Hume Series of Philosophy and Cognitive Sciences Reissues. Palo Alto, CA: Center for the Study of Language and Information, 1999).

allow us to extend the influence of present deliberation to the future. Intentions are, so to speak, the building blocks of such plans; and plans are intentions writ large.¹²

Plans themselves are typically *partial* insofar as details are added in as needed and have a *hierarchical structure* (i.e., general plans determine more particular sub-plans).¹³ On this model, intentions are guided by norms of practical rationality, such as *consistency constraints* (both internally and with an agent's belief), *means-ends coherence*, *agglomeration* (i.e., group smaller intentions into a larger intention), and *stability*.¹⁴

Intentions function as a commitment. Future-oriented intentions, in particular, are volitional commitments that “resist (to some extent) revision and reconsideration”¹⁵ and are governed by *internal norms* of rationality, which inform of the content of present and future intentions by constraining admissible options for action.¹⁶ Prior intentions and plans constitute a *background framework* for development of present and future-oriented intentions insofar as they provide *framework reasons*, which are reasons that allow an agent to consider more particular desire-belief reasons for an intention.¹⁷ Intentional actions are judged as rational from two different perspectives: a plan-constrained rationality (i.e., an *internal-ought judgment*) and a non-plan-constrained rationality (i.e., an *external-ought judgment*).¹⁸ Plan-constrained rationality is a central feature of an agent's capacity for deliberation, because the decision to do X or Y, for

¹² Bratman, *Intentions, Plans, and Practical Reason*, 8.

¹³ Bratman, *Intentions, Plans, and Practical Reason*, 29.

¹⁴ Bratman, *Intentions, Plans, and Practical Reason*, 30-32, 65-70, 72-73, 134-138; *SA* 15-25.

¹⁵ Bratman, *Intentions, Plans, and Practical Reason*, 108.

¹⁶ Bratman, *Intentions, Plans, and Practical Reason*, 109.

¹⁷ Bratman, *Intentions, Plans, and Practical Reason*, 32-35.

¹⁸ Bratman, *Intentions, Plans, and Practical Reason*, 45-46.

instance, must cohere with prior plans and intentions. Non-plan constrained rationality allows a third-party perspective to bracket an agent's prior plans and intentions in order to determine whether the action is rational insofar as it coheres with the agent's desire-belief reasons for the action.

Bratman's account not only concerns the rationality of intentional action but also the rationality of the agent herself. An agent's rationality is determined by a set of standards in which failure to meet such standards means that the agent is *criticizably irrational*.¹⁹ Judgments concerning an agent's rationality are external insofar as they can be made by a third party. Such judgements are based on standards of reasonableness, namely: "the agent's relevant habits, dispositions, and ways of arriving at decisions and actions come up to a certain level of effectiveness in their expected impact on that agent's long-term interest in getting what she wants."²⁰ If an agent wants to save money, then the agent is irrational if that agent fails to budget, for instance. Standards of reasonableness do not require the agent to be *optimally effective* in working toward long-term desire-satisfaction, but rather that they are generally effective, i.e., above a certain threshold.²¹ Moreover, practical rationality overtime is a skill of knowing *how*, which means knowing "*when* to reason (or not reason) about what."²² An agent must know when to reconsider prior intentions that provide framework reasons for present- and future-oriented in intentions.²³ Because there are resource limitations on our reasoning processes,

¹⁹ Bratman, *Intentions, Plans, and Practical Reason*, 51.

²⁰ Bratman, *Intentions, Plans, and Practical Reason*, 53.

²¹ Bratman, *Intentions, Plans, and Practical Reason*, 53.

²² Bratman, *Intentions, Plans, and Practical Reason*, 66. Bratman relies on Gilbert Ryle's distinction of knowing *that* and knowing *how*. See: Gilbert Ryle, *The Concept of Mind* (London Hutchinson and Company, 1949), chap. 2.

²³ Bratman, *Intentions, Plans, and Practical Reason*, 66-67.

we cannot deliberate about every intention or reconsider every prior intention repeatedly in trying to reach a goal. Rather, we must know when to deliberate and when to reconsider, which means that intentions function as commitments and some commitments become non-negotiable.

Stability is a central norm, with two tiers, that governs the terms of reconsideration:

1. *Reasonable stability*: “the stability of an intention or plan is *reasonable* if the associated habits of reconsideration are reasonable for the agent to have—if the expected impact of these habits on the agent’s long-term interest in getting what she (rationally) wants exceeds an appropriate threshold.”²⁴

2. *Ideal stability*: a question of whether reconsideration of a prior intention “would be recommended from our external, non-plan-constrained perspective as superior to just going ahead and non-deliberatively executing that intention.”²⁵

For reasonable stability, habits of reconsideration are subject to rational pressure in favor of increased stability, because the costs of reconsideration can be considerably high. As temporally constrained agents with limited resources, we do not have the luxury of constantly reconsidering all our intentions or plans.²⁶ Constant reconsideration would inhibit our ability to develop and carry out plans, so habits of reconsideration track when an agent has good reasons to reconsider a prior intention. Moreover, the increased stability of an individual’s intentions and plans makes them more dependable partners in shared planning projects.²⁷ For ideal stability, reconsideration is recommended only when the reconsideration results in a change in intention and the benefit of that change is greater than the costs of reconsideration, i.e., loss of time, resources, et cetera. Reasonable and ideal stability also allow for a division between reasonableness of habits and

²⁴ Bratman, *Intentions, Plans, and Practical Reason*, 72.

²⁵ Bratman, *Intentions, Plans, and Practical Reason*, 72.

²⁶ Bratman, *Intentions, Plans, and Practical Reason*, 67.

²⁷ Bratman, *Intentions, Plans, and Practical Reason*, 67.

agent rationality. An agent can rationally develop unreasonable habits of reconsideration,²⁸ and the reverse is true: it might not be rational for an agent to develop more reasonable habits of reconsideration.²⁹

Individual planning agency is the basis for shared planned agency. This reflects Bratman's broader *continuity thesis* modeled off of Paul Grice's creature construction: complex forms of agency and sociality are built from simpler forms.³⁰ Bratman's *augmented individualism* prioritizes ontological parsimony and rejects ontologically substantive accounts of shared intentions that proposed new attitudes (e.g., John Searle's *we-intention*) and new agents (e.g., Margaret Gilbert's *plural subject*).³¹ His account, however, is not reductive. Rather, his constructivism about shared agency aims to capture the phenomenon by building on pre-existing theories and concepts about individual agency (as opposed to developing new theories or concepts). Thus, Bratman's provides a set of success conditions for sharing intentions that follows from how individual agents develop and carry out individual intentions. In order for two or more agents to successfully share an intention, there are five sufficient conditions:

- A. *Intention condition*: we each have intentions that we *J*.
- B. *Belief condition*: we each believe that if the intentions of each in favor of our *J*-ing persist, we will *J* by way of those intentions, and we each must believe there is interdependence in persistence of those intentions of each in favor of our *J*-ing.
- C. *Interdependence condition*: there is interdependence in persistence of the intentions of each in favor of our *J*-ing.

²⁸ Bratman, *Intentions, Plans, and Practical Reason*, 70.

²⁹ Bratman, *Intentions, Plans, and Practical Reason*, 68.

³⁰ Bratman, *Shared Agency*, 7-9 25-26.

³¹ Bratman, *Shared Agency* 6-9; 104-106; 121-131. Bratman relies on David Lewis' distinction between qualitative and quantitative parsimony. Qualitative parsimony tracks how many different *types* of entities are proposed whereas quantitative parsimony tracks how many different *tokens* of the *types* proposed. One can be, as Bratman is, qualitatively parsimonious without committing to quantitative parsimony. See: David Lewis, "Psychophysical and Theoretical Identifications," *Australasian Journal of Philosophy* 50 (1973): 33-36.

D. *Common knowledge condition*: It is common knowledge that *A-D*.

E. *Mutual responsiveness condition*: our shared intention to *J* leads to our *J*-ing by way of public mutual responsiveness in sub-intention and action.³²

Shared intentions must reflexively interlock, and each agent must understand the other as an *intentional co-participant*.³³ These conditions generally rule out shared intentional activity involving coercion and deception.³⁴

By building from individual planning agency to shared planning agency, Bratman argues for a form of modest sociality, which captures how individuals interpersonally coordinate their plans and actions via their individual agency.³⁵ Essentially, there is an interconnected planning agency between individual agents whose intentions and agency are governed simultaneously by the norms of rationality at the individual and social level. As Bratman writes,

Conformity to social rationality norms that are central to shared intention — norms of social agglomeration, social consistency, social coherence, and social stability—will emerge from the norm-guided functioning of these interrelated attitudes of the individuals. Violation of such social norms will normally consist of a violation of associated norm of individual planning agency.³⁶

Like individual planning agency, shared intentions are subject to rational pressure, i.e., social rationality. Social rationality frames how two or more agents lay out the terms of interpersonal bargaining and/or shared deliberation, functioning as a form of mutual obligation.³⁷ Such

³² Bratman, *Shared Agency*, 152.

³³ Bratman, *Shared Agency*, 48-52.

³⁴ Bratman, *Shared Agency*, 37-38, 101-02.

³⁵ Bratman, *Shared Agency*, 29.

³⁶ Bratman, *Shared Agency*, 87.

³⁷ Bratman, *Shared Agency*, 108-113.

obligations arise from the condition of common knowledge and function as *mutual rational support*.³⁸ Mutual obligations are not moral, but are *contingently morally realizable*, which explains why many real-world cases will likely also contain moral obligations.³⁹

Bratman builds on his basic model for shared intention in order to explain how more complex social phenomena, i.e., shared deliberation and social networks, develop. Broadly, shared deliberation is different from ordinary bargaining, because the grounds for deliberation are based on shared commitments. This shared commitment is not a public convergence in individual value judgments within a group, but rather a shared intention that provides a “settled, public common ground that serves as a framework for relevant social thought and action,”⁴⁰ particularly temporally extended thought and action. Mirroring the role of personal policies for deliberation for individual planning agency, these shared intentions often result in general policies about weights that govern the terms of deliberation. Broadly, a shared policy is a shared commitment to give weight to *R* in relevant shared deliberation, and this shared policy emerges from public, interlocking intentions from the agents involved. Agents can hold shared commitments for different reasons and such commitments need not accord with their personal value judgments. This partial sociality allows agents from different backgrounds to engage in shared deliberation. Shared policies function like the *standpoint of the group*,⁴¹ not a convergence of individual commitments. Bratman’s account of shared deliberation provides additional support for the role of social rationality in shared intentions. Sharing intentions starts

³⁸ Bratman, *Shared Agency*, 110-111.

³⁹ Bratman, *Shared Agency*, 113.

⁴⁰ Bratman, *Shared Agency*, 137.

⁴¹ Bratman, *Shared Agency*, 142. Bratman’s use of standpoint here should not be confused with standpoint theory in feminist epistemology.

with individual agents whose intentional content reflects norms of individual rationality. This gives way to norms of social rationality that agents individually accept. Finally, individuals can develop a shared policy about the norms of social rationality, providing grounds for shared deliberation between agents. This connects with Bratman's account of how large groups function, by framing temporally extended groups in terms of quasi-Lockean social ties.⁴² A larger group need not necessarily hold a shared intention, but the ties between participating agents are more substantive than merely an aggregation of small groups.⁴³ For instance, a large-scale, intergenerational research project will have a social network that contains a number of overlapping instances of modest sociality. However, larger groups can also engage in shared deliberation. In this way, Bratman builds from small-scale to large-scale forms of modest sociality by building in complexity through the reiteration of different kinds of planning agency.

ii. Margaret Gilbert: Joint Commitments

Gilbert's account of *joint commitments* offers a normative analysis that posits a *plural subject* as the subject of the shared intention.⁴⁴ Gilbert broadly rejects *singularist* accounts of social phenomenon (e.g., Bratman and Kutz) as they prioritize a game-theory model of individuals and their behavior, leaving no room for *cogitamus*, or a we-mode intentional state.⁴⁵

⁴² Bratman, *Shared Agency*, 98.

⁴³ Bratman, *Shared Agency*, 99.

⁴⁴ Gilbert, *Joint Commitment*.

⁴⁵ Gilbert, *Joint Commitment*, 4, 94-128. Singularist accounts posit interlocking personal intentions, i.e., Bratman, or contributory individual intentions, i.e., Kutz.

Gilbert regards joint commitments as foundational to understanding rights and obligations, particularly those not captured well by the framework of *moralism*.⁴⁶

For Gilbert, the term ‘commitment’ refers to “commitments of the will,” in which a commitment is a binding decision to engage in an activity unless the commitment is altered or rescinded.⁴⁷ Commitments have a rational structure insofar as the agent who makes the commitment “has sufficient reason to act in accordance with it.”⁴⁸ The ‘ought’ of a commitment is not moral but rather rational.⁴⁹ There are both *personal* and *joint commitments*. A personal commitment is a commitment of person Z in which Z, as the sole subject of the commitment, can unilaterally both bring about and rescind it. If Z decides to go drink a glass of water, Z can rescind this commitment or rebuke themselves for violating it by failing to act. By contrast, a joint commitment occurs when two or more people make a commitment *to ϕ as a body*, where ϕ constitutes the relevant verb.⁵⁰ The subject of a joint commitment is a newly created *plural subject*, i.e., *we* (or the plural *they*) rather than *I*. A plural subject is not the accumulation of the personal commitments of members of the commitment nor completely reducible to individual members.⁵¹ Rather, a joint commitment creates a plural subject insofar as all individual participants are committing to *ϕ -ing* as a unified body, i.e., such a commitment is non-

⁴⁶ Gilbert, *Joint Commitment*, 5-9, 85-89, 108-113. *Moralism* is a framework in which talk of rights or obligations is understood as moral in nature, be it questions of moral worth or requirement.

⁴⁷ Gilbert, *Joint Commitment*, 83-85.

⁴⁸ Gilbert, *Joint Commitment*, 84.

⁴⁹ Gilbert, *Joint Commitment*, 84.

⁵⁰ Gilbert, *Joint Commitment*, 7.

⁵¹ Gilbert, *Joint Commitment*, 8-9.

distributive.⁵² A joint commitment can subsequently generate *individual commitments*, which are the individual roles or actions required to ϕ as a unified body.

Answerability is a key feature of personal and joint commitments. If Z makes a personal commitment to drink more water, Z can rebuke themselves for failing to do so. In contrast, if Z and X make a joint commitment to meet for lunch, Z and X are now answerable to each other. Rescinding the commitment requires the permission of all committed parties, because entering into the commitment generates a set of obligations and rights for all participants. If an individual violates a joint commitment, other participants are justified in normative rebuke. Such commitments are binding insofar as participants owe one another the actions required to fulfill the terms of the commitment.⁵³ These are *directed obligations* that give participants mutual claims against one another and thus, by extension, the standing to demand fulfillment or rebuke violation of the commitment.⁵⁴ The answerability feature ensures a greater stability over time because it binds participants normatively and offers a stronger foundation for bargaining, negotiating, and coordination in maintaining or carrying out the commitment.⁵⁵

Joint commitment is not an agreement between two or more parties to ϕ together, although explicit agreements are present in many joint commitments. A joint commitment arises when two or more agents express a *personal readiness* to enter into the commitment. Personal

⁵² Gilbert, *Joint Commitment*, 88-93, 114-120. The plural subject is the subject of the intention but does not itself have subjectivity or consciousness. For theorists like Bratman or John Searle, the individual is the subject of the intention. Searle, in particular, argues that an intention must exist within a mind, brain, and/or within consciousness. See: John Searle, "Collective Intentions and Actions," in *Intentions in Communication*, ed. Philip R. Cohen, Jerry Morgan, and Martha Pollack (Cambridge: MIT Press, 1990), 401-415. In this respect, Gilbert's plural subject is more similar metaphysically to accounts of group agency, such as those offered by List and Pettit (2011) and Tuomela (2013).

⁵³ Gilbert, *Joint Commitment*, 109.

⁵⁴ Gilbert, *Joint Commitment*, 109-113. Gilbert refers to this as *the obligation criterion*.

⁵⁵ Gilbert, *Joint Commitment*, 123-124.

readiness is an expression of willingness by all involved parties to enter into the commitment.⁵⁶

All potential participating agents are in a state of personal *quasi-readiness* and this makes them *jointly ready*.⁵⁷ Such expressions can range from the non-verbal development of a shared routine over time (e.g., walking to lunch together each day), an immediate expression through mutual engagement (e.g., quarreling), to an explicit agreement between the parties. This readiness is further mediated by a *condition of common knowledge* in which individual expressions of readiness are common knowledge to all participating parties and this knowledge is *open* to all parties.⁵⁸ Common knowledge is possible when agents are “normal human beings”⁵⁹ with “similar conceptual equipment”⁶⁰ who perceive each other and perceive these shared traits. As Gilbert writes:

There is *common knowledge that p* among certain parties if and only if the parties *notice* that the fact *that p* is *open* with respect to all of themselves. As I define it, *openness* involves, roughly, many levels of potential knowledge of one another’s knowledge.⁶¹

Moreover, there are two types of common knowledge: *individual* and *population*.⁶² Individual common knowledge occurs when all parties know of and, in fact, know each other, e.g., two friends jointly committing to take a walk together. Population common knowledge exists

⁵⁶ Gilbert, *Joint Commitment*, 47-48.

⁵⁷ Margaret Gilbert, *On Social Facts* (Princeton: Princeton University Press, 1989), 184-185.

⁵⁸ Gilbert, *Joint Commitment*, 29, 43, 51.

⁵⁹ Gilbert, *On Social Facts*, 188. “In particular they have normal perceptual organs functioning normally, and they have normal reasoning capacities.”

⁶⁰ Gilbert, *On Social Facts*, 188. “In particular they have the concepts they need in order to fulfill the other conditions.”

⁶¹ Gilbert, *Joint Commitment*, 43. For a more complete discussion of *openness* and its basis in set theory, see: Gilbert, *On Social Facts*, 188-191.

⁶² Gilbert, *Joint Commitment*, 51-52.

between *distanced* populations wherein individuals do not know or know of each other but can still have knowledge of each other's intentional states if such intentions are expressing openly.⁶³

Gilbert builds on the basic structure of joint commitment in order to capture larger questions of shared values, coercive law, social groups, and, ultimately, political obligation. A joint commitment is a precursor for joint actions, but it is also a precursor for collective belief, collective emotions, shared values, laws, and social norms.⁶⁴ In cases of shared belief or value, members jointly commit to believing that *p* if and only if all members commit to upholding that belief as part of the plural subject. Joint belief does not require that all members personally believe that *p*, but rather that a person must commit to upholding that belief as part of the plural subject.⁶⁵ Failure to do so justifies rebuke from other members. Moreover, the presence of coercion or duress does not prevent one from entering into a joint commitment. One need only to express a readiness to enter into the commitment, even if that readiness is achieved coercively and/or does not reflect one's individual intentions, beliefs, or values. Membership in a plural subject necessarily obligates one to act in accordance with it, even if the commitment is immoral or brought about coercively.

From this foundation, Gilbert develops a plural subject theory of political society and political obligation, focusing on the terms of group membership, i.e., *our* country. Broadly, joint commitments capture how political societies form (via smaller social groups) and how political obligation functions for members of a political society, justifying the presence of coercive law and punishment. Gilbert rejects social contract theory, i.e., *actual contract theory*, as the basis of

⁶³ Gilbert, *Joint Commitment*, 51-52.

⁶⁴ Gilbert, *Joint Commitment*, 131-162, 181-258.

⁶⁵ Gilbert, *Joint Commitment*, 131-162, 193-194.

political membership and obligation, arguing that an agreement model is too narrow to realistically capture political society as a social group with sometimes fuzzy boundaries for membership. Plural subject theory posits commitments as prior to agreements (of which contracts are one form). On this model, plural subjects constitute social groups that generally involve the perception of unity insofar as the groups' existence has some intentional basis (if only in core or founding members) and member identification with the group, i.e., the use of *us*, *our*, or *we*. Plural subjects can be temporally extended, depending on the nature of the group. In limited contexts, plural subjects can be defined by their associative bonds, such as the joint commitment to a set of shared values. Agreements are sufficient but not necessary to establish plural subjects, and not all plural subjects necessarily have a goal or an aim, i.e., a family.

Although with social groups, societies can constitute plural subjects. Gilbert's reference to John Rawls' claim that a political society is a "social union of social unions" is noteworthy here.⁶⁶ Smaller plural subjects can constitute the membership of larger, inclusive plural subjects.⁶⁷ In hierarchical plural subjects, designated members are able to make *derived* joint commitments on behalf of the members of the initial commitment, which renders non-designated members subject to those derived commitments, even if they do not know they are.⁶⁸ On a large scale, this can result in the development of three kinds of social rules: *governing rules*, *personal rules* (i.e., rules stipulating some person or body's ruling capacity), and *rules of governance or constitutional rule*.⁶⁹ From the condition of (imperfect) population common knowledge, it

⁶⁶ Gilbert, *A Theory of Political Obligation* (Oxford: Oxford University Press, 2006), 17. See: John Rawls, *A Theory of Justice*, rev. ed. (Cambridge, MA: Belknap Press of Harvard University Press, 1999), 527.

⁶⁷ Gilbert, *A Theory of Political Obligation*, 173-174.

⁶⁸ Gilbert, *Joint Commitment*, 352-353.

⁶⁹ Gilbert, *Joint Commitment*, chap. 9 (esp. 204-214), 288.

follows that even if there is impersonality and anonymity amongst most members, there can still be large-scale expressions of readiness that are engaged in and knowable to most members of that population.⁷⁰ Looking toward the *Crito*,⁷¹ Gilbert argues that members of political societies are jointly committed to upholding political institutions (and the social rules which comprise them) and must accept punishment for failing to do so, even if the political society in question is fundamentally immoral.⁷² If one does not except a particular rule, then the burden is on the individual to try to persuade that society to rescind the particular joint commitment that has created it, but until this commitment is jointly rescind, the individual is obligated to follow it or accept punishment. However, since obligation as owing is not strictly moral, individuals can still have particular moral obligations that conflict with or override the obligations of joint commitments.

Stepping back, Gilbert's theories of joint commitment and plural subjecthood establish a foundation for shared intention that captures everything from two people taking a walk together to the basis for large-scale political society and political obligation. This model of shared intention is universal insofar as it applies to all "normal" rational agents,⁷³ regardless of more particular features of their agency. Moreover, the account of political obligation extends beyond the standard focus on liberal democracies and describes all types of political societies.⁷⁴

⁷⁰ Gilbert, *A Theory of Political Obligation*, 174-179.

⁷¹ Plato, *Crito*, in *Five Dialogues: Euthyphro, Apology, Crito, Menno, Phaedo*, ed. John Cooper, trans. G.M.A. Grube (Indianapolis, IN: Hackett Publishing Company, 2002), 45-57.

⁷² Gilbert, *A Theory of Political Obligation*, 3-4, 290-293.

⁷³ Gilbert, *A Theory of Political Obligation*, 174. Here Gilbert includes a developmental claim about age. Elsewhere, Gilbert has defined the normalcy of agents in terms of capacities for reasoning and the presence and function of perceptual organs (see: n64-65).

⁷⁴ Gilbert, *A Theory of Political Obligation*, 25.

iii. Christopher Kutz: Participatory Intentions

Kutz's minimalist conception of a *participatory intention* aims to capture the potential range and complexity of collective actions.⁷⁵ By emphasizing generality, reducibility, and functionalism,⁷⁶ his account of joint action is metaphysically and psychologically parsimonious. He begins from an account of agency that treats any collective action as reducible to the individual intentions of individual agents. In this sense, participatory intentions are one form of what Margaret Gilbert calls (*personal*) *contributory intention*.⁷⁷ For Kutz, such intentions are a sufficient condition for collective action. One need not postulate shared intentions with more substantive success conditions, plural subjects, or group agents with irreducible group attitudes. Rather, an account of collective action need only understand how individuals develop personal intentions to coordinate with others toward a collective goal.

Participatory intentions rely on Kutz's broader account of relational and positional accountability, which rejects ethically solipsistic forms of individual accountability.⁷⁸ Ethical solipsism⁷⁹ follows from the doctrine of methodological solipsism in philosophy of psychology and claims that: "the intentional content of an agent's mental states must be consistent with the possibility of metaphysical solipsism, which is to say, independent of facts external to the

⁷⁵ Kutz, *Complicity*, 66-112; Christopher Kutz, "Acting Together," *Philosophy and Phenomenological Research* 61, no. 1 (2000): 1-31.

⁷⁶ Kutz, *Complicity*, 68-74. Generality accounts for differences in: number of agents, task intricacy, cooperative spirit, agent autonomy, and egalitarianism. Reducibility is a concern about the causal explanation for joint actions (i.e., that group actions can be rephrased in terms of individual contributions to the group act). Functionalism offers a model for individual intentions that relies on a general theory of human rationality in which individual behavior is interpretable to a third party insofar as all intentional states are the result of a reason for that intention.

⁷⁷ Gilbert, *Joint Commitment*, 103.

⁷⁸ Kutz, *Complicity*, 10, chapter 2.

⁷⁹ Kutz, *Complicity*, 4. Kutz also refers to this as *evaluative solipsism*.

agent.”⁸⁰ Kutz rejects ethical solipsism for two reasons. First, it ignores the relationships and roles among agents, victims, and bystanders, which Kutz argues are conceptually central to understanding a warranted response for reasons of conduct, consequences, or character.⁸¹ Second, ethical solipsism not only ignores the importance of such relations but is also causally solipsistic because it reduces causal responsibility to facts about the individual agent, ignoring the possibility of joint causation between multiple agents in collective action contexts (i.e., the possibility of complicity, in particular).

In contrast, Kutz’s goal is to provide a model for individual accountability that avoids the pitfalls of ethical solipsism. Participatory intentions thus provide a way to show how individuals are accountable for individual harms in the case of collective actions. In the case of a collective action, individual agents act with the intention to contribute to the collective end, or goal, which makes the content of such an intention irreducibly shared. A participatory intention has two conditions of satisfaction: (1) the individual role, and (2) the collective end. Each agent understands their individual role as contributing to a collective end and recognizes themselves as doing their part to contribute to that end in concert with others. On a phenomenological level,⁸² individual agents must regard themselves as contributing to the collective end. This *internalist* requirement tracks how individuals conceptualize their own agency within a group-oriented context.⁸³ In this way, participatory intentions challenge a classic individualistic model for rational choice theory, because individual actors often have some knowledge of others’

⁸⁰ Kutz, *Complicity*, 4n13.

⁸¹ Kutz, *Complicity*, 24-26.

⁸² Kutz, *Complicity*, 5. Kutz initially defines his use of ‘phenomenology’ as: the sense we have of ourselves as individual agents that “emerges from the experience of making changes in our environment.” I assume this is broadly how he uses the word throughout the text.

⁸³ Kutz, “Acting Together,” 17.

preferences and can use this in their individual deliberation to opt for the collectively rational outcome.⁸⁴ In collective actions, the individualistic strategic reasoning of game theory does not work because agents' choices can be interdependent. Rather than acting toward one's individually preferred end, agents can want to simply match their preferences, rendering the preference indeterminate prior to matching. It is only when deliberating about the preferable course of action for the collective (rather than the individual) that agents can develop a shared preference that informs the content of their intentions. For Kutz, this interpersonally sensitive deliberation alters, what Susan Hurley refers to as, "the conceptual unit of agency,"⁸⁵ insofar as agents come to regard one another as intentionally cooperative participants in a joint action. In this way, participatory intentions are "merely a species of ordinary, instrumental intentions, differentiated by the group-oriented context."⁸⁶

Thus, an action is collective when all individuals have an intention to do their part in bringing about the collective end, wherein an individual's part is defined as "the task [one] ought to perform if [the collective is] to be successful in realizing a shared goal."⁸⁷ A participatory intention is then the goal to accomplish one's individual role insofar as it contributes to the collective end.⁸⁸ The condition of extensional overlap defines when agents, in fact, share a goal or end by stipulating the boundaries of a joint enterprise: "Agents' intentions overlap – they share goals – when the collective end component of their participatory intentions refers to the

⁸⁴ Kutz, "Acting Together," 9.

⁸⁵ Kutz, "Acting Together," 9; Kutz, *Complicity*, 275n29. See: Susan Hurley, *Natural Reasons* (New York: Oxford University Press, 1988), 145-48.

⁸⁶ Kutz, "Acting Together," 12.

⁸⁷ Kutz, "Acting Together," 10.

⁸⁸ Kutz, "Acting Together," 10.

same activity or outcome and when there is a nonempty intersection of the sets of states of affairs satisfying those collective ends.”⁸⁹ Such overlap need not be perfect nor substantive.

Kutz argues that participatory intentions are characterized by *mutual openness* and *strategic responsiveness*.⁹⁰ Mutual openness is weaker than David Lewis’ common knowledge condition,⁹¹ used frequently by a number of shared intention accounts. Rather than argue that all participants must know that *p* and know that each knows that *p* (and so forth), Kutz argues that mutual openness, as a version of Sperber’s and Wilson’s concept of *mutual manifestness*, captures a wider range of potential joint actions.⁹² As a language claim, mutual manifestness appears when agents have shared cognitive background that is not the explicit content of the beliefs but can help disambiguate ambiguous utterances.⁹³ Mutual openness occurs when agents have “dispositions favorable” to mutual manifestness.⁹⁴ Strategic responsiveness is a feature of an agent’s participatory intention, tracking how an agent’s intentions develop in relation to a collective goal. This occurs when an agent’s intentions are “sensitive to their beliefs or predications about what others intend to do.”⁹⁵ Conditions of mutual openness and strategic responsiveness ensure that agents are acting cooperatively insofar as they understand the shared end and their reasons for action take that end into account. However, as long as the condition of

⁸⁹ Kutz, *Complicity*, 94.

⁹⁰ Kutz, *Complicity*, 76-77.

⁹¹ Kutz, *Complicity*, 274n22. See: David Lewis, *Convention* (Cambridge, MA: Harvard University Press, 1968), 52-60.

⁹² Kutz, *Complicity*, 274n22. See: Dan Sperber and Dierdre Wilson, *Relevance* (Cambridge, MA: Harvard Press, 1986), 38-45.

⁹³ Kutz, *Complicity*, 274n22.

⁹⁴ Kutz, *Complicity*, 274n22.

⁹⁵ Kutz, *Complicity*, 76.

extensional overlap holds, the conditions of mutual openness and strategic responsiveness are not necessary for an action to be jointly intentional.⁹⁶ Participatory intentions merely require that I have an intention to do my part in enacting the collective goal with a loose awareness that other participants are doing the same.

Participatory intentions can be further understood in terms of *executive* and *subsidiary* intentions.⁹⁷ An executive intention is a second-order participatory intention that plays a determining role in the content of first-order, subsidiary intentions. Subsidiary intentions can be either individual or participatory, but their content fulfills an instrumental role in relation to the collective end. It is not required that all participating agents hold both intentions. Rather, some individuals might hold an executive intention (or even both types of intention) while many others only hold subsidiary intentions. Since individuals can hold just a subsidiary intention, even those marginal contributions can intentionally contribute to a collective end without holding the executive intention. Kutz further distinguishes between intentional and intended action:

“individuals can intentionally contribute to a collective end even though they do not intend the realization of that end.”⁹⁸ Agents can thus be alienated from the end to which they contribute. To use Kutz’s example, a pacifist might take a job at a nuclear weapons plants because that person lacks other employment options, despite not intending the realization of a nuclear weapon. Moreover, the intent to participate does not require that the agent positively contribute to the collective end, i.e., the agent can actually hinder it, nor do participatory intentions require that

⁹⁶ Kutz, *Complicity*, 92.

⁹⁷ Kutz, “Acting Together,” 14.

⁹⁸ Kutz, “Acting Together,” 25.

agents believe in the potential success of the collective action for joint intentionality to be present.⁹⁹

Since participatory intentions are necessarily found within individual participants and do not require participants to develop strong beliefs or expectations about the intentions of other participants, Kutz regards claims about group intention to be too narrow to capture the potential range of collective actions.¹⁰⁰ An agent can intend to do their part in *J*-ing without the intentional structure of “I intend that we *J*” or “We intend to *J*.” Kutz rejects Bratman’s and Gilbert’s accounts because each requires agents to act in relation to “positive expectations about one another’s plans.”¹⁰¹ One can contribute to the collective action without assuming the executive intention of the action, i.e., the collective end. Since participatory intentions are individual in this way, one can hold a participatory intention to begin a collective action without knowing whether others will join in.¹⁰² Broadly, Kutz’s minimalist account of the sufficient condition for collective action is thus the following:

All collective action, hierarchical and non-hierarchical, pre-programmed and dynamic, planned and spontaneous, admits of one common analysis: a set of individuals jointly *G* when the members of that set intentionally contribute to *G*’s occurrence by doing their particular parts, and their conceptions of *G* sufficiently and actually overlap.¹⁰³

This account allows for collective action to occur even when the agents are loosely linked and have little information about one another’s mental states, i.e., intentions, beliefs, desires, so long

⁹⁹ Kutz, *Complicity*, 84, 92.

¹⁰⁰ Kutz, *Complicity*, 96-103.

¹⁰¹ Kutz, *Complicity*, 96.

¹⁰² Kutz, *Complicity*, 92.

¹⁰³ Kutz, “Acting Together,” 27.

as each individual agent has an awareness that they are acting in concert with others and their individual role contributes to the collective end, however successfully.

Section III. The Methodology for Shared Agency

Each of these three accounts provides a model for how joint actions occur by way of shared or participatory intentions. In this way, each account not only offers a set of success conditions for sharing an intention, but also success conditions for the joint action that follows from it. That is, these models do not merely aim to examine how people share intentions but also by extension explain what a joint action is. In this section, I revisit and expand upon the discussion of idealizations from Chapter I, explaining how each of the above models relies on a similar set of idealizations: (1) the idealization of agent-neutrality; (2) an idealized epistemology (and cognitive sphere); and (3) an idealized social ontology that ignores the existing of forms of *ontic injustice*.¹⁰⁴

In *As If*, Appiah argues that idealizations are a fundamental feature of how we theorize about ourselves and the world around us. Drawing on the work of Hans Vaihinger,¹⁰⁵ Appiah argues that idealizations function as an as-if claim about the world that is productive for both understanding the world and allowing us to navigate this world more easily. To employ an as-if claim, also known as an idealization, a theorist knowing relies on a falsehood in studying a phenomenon. To use Appiah's example, Kant's account of rational agency requires us to act *as if* we have free will, despite the fact that we are simultaneously aware, at least theoretically, of

¹⁰⁴ Katharine Jenkins, "Ontic Injustice," *Journal of the American Philosophical Association* (forthcoming).

¹⁰⁵ Hans Vaihinger. *The Philosophy of "As If": A System of the Theoretical, Practical, and Religious Fictions of Mankind*, trans. C.K. Ogden (New York: Harcourt, Brace, and Co., 1924).

being governed by deterministic laws.¹⁰⁶ In this way, idealizations are ‘useful untruths.’¹⁰⁷

Appiah’s analysis of these useful untruths relies heavily on Vaihinger’s theory and taxonomy of idealizations. Vaihinger’s account of idealization is broadly pragmatist:

It must be remembered that the object of the world of ideas as a whole is not the portrayal of reality — this would be an utterly impossible task — but rather to provide an instrument for finding our way about more easily in the world.¹⁰⁸

However, the core of Vaihinger’s account still relies on a distinction between what is useful to believe about the world and what is true about the world – thus his understanding of idealizations as useful untruths, or a form of fiction. Vaihinger’s taxonomy of fictions argues that there are ‘real fictions’ (i.e., a claim that is false and contradictory) and ‘semi-fictions’ (i.e., a claim that is false but not contradictory), which both fall under ‘scientific fictions,’ noting, in addition, a variety of other forms: abstractive fictions (i.e., abstractions), average fictions (i.e., the appeal to averages) schematic fictions (which include the fiction of the simple case), and heuristic fictions (i.e., former hypotheses that are still scientifically useful).¹⁰⁹ Appiah also notes a more contemporary division in forms of fictions: Galilean idealizations (i.e., idealization by abstraction) and Aristotelian idealizations (i.e., idealization by approximation).¹¹⁰ Galilean idealizations are those that introduce “distorting simplifications”¹¹¹ for reasons of computational

¹⁰⁶ Appiah, *As If*, 3.

¹⁰⁷ Appiah, *As If*, 3.

¹⁰⁸ Appiah, *As If*, 5. Appiah defines Vaihinger as pragmatist, while Vaihinger himself notes that there is a difference between his factionalism and pragmatism. However, Vaihinger does argue that the end result of a fictionalist or a pragmatist approach are effectively the same. See: Vaihinger, *The Philosophy of ‘As If,’* viii.

¹⁰⁹ Appiah, *As If*, 9-10.

¹¹⁰ Appiah, *As If*, 10-11. Here Appiah’s analysis draws on: Ernan McMullin, “Galilean Idealization,” *Studies in History and Philosophy of Science*, Part A, no. 3 (1985): 247-273. McMullin’s article highlights the range of idealizations used in the sciences, starting from mathematical idealizations.

¹¹¹ Appiah, *As If*, 10.

tractability, such as understanding motion by postulating a frictionless plane or starting from the assumption that agents are perfectly rational; whereas Aristotelian idealizations work by simplifying the phenomenon under investigation by removing complicating factors that are seen as negligible (e.g., the color of a ball has no bearing on its movement and thus can be removed in studying its motion).

What is valuable about Vaihinger's taxonomy is that rather than make a hard distinction between idealizations and abstractions (as Onora O'Neill does), it shows that abstraction is a form of idealization because abstraction by its very nature postulates a fiction.¹¹² It offers, instead, a broader definition for what constitutes an idealization and is able to show how idealizations are used in different contexts for different purposes. Thus, for Vaihinger and, by extension, Appiah, theorizing necessarily relies on some form of fiction about the world and the form of fiction, or idealization, that a theorist is justified in relying on depends on the broader context. To avoid idealization is to attempt to embrace the radical complexity of the world all at once, which is impossible, because, as Appiah rightly notes, "The complexities exceed *our* cognitive capacity to encompass *them*, and that is as much a fact about *us* as about *them*." Thus, contrary to those who dismiss the use of idealizations outright, it is the complexity of the world that requires theorists to engage in some form of idealization.

However, the claim that idealizations are necessary does not entail the claim that all idealizations are useful, nor does it follow that all idealizations are useful for any purpose. Rather, the justified use of idealization depends on what kinds of model of the world one is

¹¹² Onora O'Neill, "Abstraction, Idealization, and Ideology in Ethics," in *Moral Philosophy and Contemporary Problems*, ed. J.D.G. Evans (Cambridge, UK: Cambridge University Press, 1987), 55-69; "Ethical Reasoning and Ideological Pluralism," *Ethics* 98, no. 4 (July 1988): 705-722.

trying to develop and the explanatory power of the model.¹¹³ For this reason, Vaihinger argues that we ought to oppose the use of bad idealizations just as one would oppose a bad hypothesis.¹¹⁴ Idealizations are thus good or bad depending on whether they are useful for the purpose in question. For this reason, I take the distinction between good and bad idealizations to be about justification.

Good idealizations are those that are justified by their use in a particular context insofar as they successfully fulfill one of two possible aims. Following Vaihinger, the first aim in using idealizations is to better understand the world and ourselves. For example, Newton's laws of physics rely on an idealization of objects and, in the case of his first law, objects are idealized insofar as their capacity for motion is understood independently of the presence of force (which, in actuality, is always present in some capacity). In this way, Newton's laws rely on a useful idealization because, despite starting from an untruth, they offer one model for understanding the world that is practically useful. What is noteworthy here is that the model of Newtonian mechanics can exist in contradiction (or, at least as incompatible) with the model of Einsteinian mechanics, because they offer two different levels of explanatory power. Even though Newtonian mechanics is currently understood to be incompatible with Einsteinian mechanics (and, if the latter is taken to be a more accurate explanation of reality, a kind of fiction in itself), Newtonian mechanics are still explanatorily useful in many contexts.¹¹⁵

¹¹³ Appiah, *As If*, 12-13. Appiah notes that theorists can develop different and contradictory models to study the same phenomenon. The value of each model is different because they do not serve the same purpose. Appiah likens this to David Lewis' discussion of internal maps: sometimes we have inconsistent internal maps of the same street because we use the street for different purposes. See: David Lewis, "Logic for Equivocators," *Noûs* 16, no. 3 (1982): 431-442.

¹¹⁴ Appiah, *As If*, 182n29. Vaihinger writes, "We shall indeed have at every step to oppose bad fictions, just as formerly bad hypotheses were opposed." See: Vaihinger, *The Philosophy of "As If,"* 45.

¹¹⁵ Appiah, *As If*, 27, quoting Vaihinger, *The Philosophy of "As If,"* 5.

The second aim in using idealizations, however, is not about understanding the world but instead about “managing reality.”¹¹⁶ That is, the explanatory power of a model allows us to understand phenomena and, by extension, control or predicate the behavior of the phenomenon in question. For example, Newtonian mechanics allows us to construct bridges (without any need to appeal to Einsteinian mechanics), and social sciences like psychology enable us to better predicate human behavior and manage it (e.g., our capacity to manage mental illness depends on how explanatory useful our theories are). The extent to which an idealization fulfills these aims is the basis by which we judge the idealizations in question as good or bad, as justified or not.

With this framework in mind, I now return to the three models for shared intentions (and, by extension, joint action) to highlight the idealizations used in each case. My analysis here is not intended to be exhaustive but rather to focus on how each model uses idealizations by showing (what I take to be) the most prominent idealizations used by each model. My focus on different idealizations in each case is not to suggest that the three models rely on completely different idealizations, but rather to highlight the range of idealizations used within research on shared intentions. For example, all three models rely on idealizations of agent-neutrality – an idealization that relies on abstract human agents – but do so in slightly different ways. Alternatively, only two of the three models rely heavily on simplified cases, a form of idealization, under the claim that a simple model for small-scale shared intentions can be extended to include more complex cases. By focusing on what idealizations are employed in these three models, I suggest (but do not argue for here) that many other models throughout the analytic literature on shared intentions rely on similar idealizations. Thus, whether the

¹¹⁶ Appiah, *As If*, 27, quoting Vaihinger, *The Philosophy of “As If,”* 5.

idealizations used in these three are viable likely has import for determining whether similar idealizations employed by many other models for shared (or group) intentions are also viable.

i. Bratman: Creature Construction

Bratman's model for shared intentions relies on a few key idealizations, which he captures well when acknowledging the potential limitations of his account:

The limitation is that my focus will be primarily on the shared intentional activities of small, adult groups in the absence of asymmetric authority relations within those groups, and in which the individuals who are participants remain constant over time. Further, I will bracket complexities introduced by the inclusion of the group within a specific legal institution such as marriage, or incorporation. My interest will be primarily with duets and quartets rather than symphony orchestras with conductors, with small teams of builders rather than large and hierarchical construction companies, with small group discussion rather than deliberations in the US Senate, and with friendship and love rather than legally constituted marriage.¹¹⁷

Bratman argues that his focus on small-scale shared agency does not aim to simply exclude more complex cases. Rather, by developing a model for shared intentions through these small-scale, simplified cases, it is possible to extend this model to include more complex cases as well, which he does in his analysis of shared deliberation and shared policies. One major feature of his methodology is 'creature construction,' which argues that simple forms of agency are the building blocks to more complex forms of agency.¹¹⁸ Bratman's creature construction is his first form of idealization, because the foundation of this construction is not about how planning agency actually emerges (i.e., evolutionarily and historically) but functions as a necessary fiction, as "... such a hypothetical series of constructed creatures can help us understand

¹¹⁷ Bratman, *Shared Agency*, 7.

¹¹⁸ Bratman, *Shared Agency*, 113.

complex elements of our actual planning agency...”¹¹⁹ As Appiah similarly argues, Bratman takes creature construction as a useful untruth that provides us with an enhanced understanding of how things actually are. Because Bratman’s other idealizations follow from this method of constructivism, one might argue that this functions as a second-order idealization from which his first-order idealizations follow. Thus, whether Bratman’s idealizations are viable is a two-part question. First, there is a question of whether his second-order idealization is viable and what bearing this has on his first-order idealizations. Secondly, are these idealizations viable for his small-scale model and, if so, are they also viable in the case of extension into further complexity? Answering these questions requires understanding the nature and potential value of his first-order idealizations. I focus on two interrelated sets of first-order idealizations: (1) the simplified case; and (2) the pre-social individual.

First, Bratman idealizes by focusing on highly simplified cases that lack a formal hierarchy as well as informal forms of hierarchy that might arise in cases of coercion or unjustified social hierarchies. One reason for this kind of idealization based off the simple case, as Appiah notes, is to build a model (or models) by first trying to isolate the phenomenon from other related or overlapping phenomenon.¹²⁰ The attempt to isolate a phenomenon is a useful untruth, because while the phenomenon never actually functions in isolation, our idealized version of it provides a better understanding of how the phenomenon works under actual (i.e., not isolated) conditions compared to a model that attempted to account for all possible complexity from the beginning. In fact, the latter model would be an impossibility, because it exceeds our own capacity to cognize and account for all possible complexity. In this way,

¹¹⁹ Bratman, *Shared Agency*, 25-26.

¹²⁰ Gilbert also uses the idealization of the simple case but does not include the same additional idealizations (i.e., lack of coercion, hierarchy, et cetera).

idealizations from the simple case are akin to idealizations in Newtonian mechanics: by understanding how a phenomenon works under isolated conditions, we better understand how it works in non-isolated conditions. Further, by extending our simplified case and adjusting as needed, we can see how this particular phenomenon interacts with others. In other words, if we start from Bratman's simple cases, then one can add in particular complicating conditions one by one. If we have a model for how shared intentions work prior to the introduction of formal hierarchies, then adding in formal hierarchies later not only provides us with a better understanding of shared intentions but also of how formal hierarchies function. As Bratman himself argues, adding in features like deception or coercion shows that it is likely the case that shared intentional activities are not the same thing as shared cooperative activities.¹²¹ In this way, we can build highly complex models through simplified models, which is the cornerstone of Bratman's constructivism.

Besides starting from the simple case, Bratman's model for small-scale shared agency idealizes in another noteworthy way. Shared agency is a form of modest sociality, because it retains Bratman's model for individual planning agency while adding minimal modifications that incorporate the fact that our individual planning agency does not, in fact, exist in isolation but rather in a social world with many others. Of course, this scaling-up from the individual is also taken as a useful untruth, as Bratman himself acknowledges:

This is not to say that in the course of our actual lives we ourselves make a transition from nonsocial to social creatures. Creature construction is not a story of actual human development, and it can recognize that human lives are embedded in the social from the start... What we are after is not a story of actual human development but an understanding of the conceptual, metaphysical, and normative deep structure of our sociality.¹²²

¹²¹ Bratman, *Shared Agency*, 38.

¹²² Bratman, *Shared Agency*, 30.

That is, we do not genuinely believe that individuals are pre-social, but here it is taken as a useful untruth in order to better understand how our social world actually operates. This understanding is not, as both Appiah and Vaihinger note, about the truth of the matter but rather a balance between accuracy and computational ease. In this way, the truth of the matter requires accuracy without ease, which is just another way of saying that it would be computationally intractable.

Finally, it is worth noting that Bratman's model is not merely about understanding shared intentions and joint actions but also about managing them. Shared planning agency requires a certain level of adherence to the norms of social rationality. While violating that norms of social rationality often also constitutes a violate of norms of individual rationality, this is not necessarily the case, creating a gap between the two sets of norms (or, at least, the possibility thereof). This is why the process of shared deliberation can develop a shared policy about adherence to the norms of social rationality, namely what norms we give weight to in a deliberative process. In this way, Bratman's model also provides a basis for managing forms of shared agency, particularly in relation to possible extensions for his small-scale model.

ii. Gilbert: Personal Readiness

Gilbert's model for shared intentions includes a necessary normativity that offers a slightly different set of idealizations. Gilbert's overall methodology is similar to Bratman's insofar as she begins with simple cases, building in complexity step-by-step.¹²³ Gilbert's idealizations differ from Bratman's insofar as her method of building in complexity is not a form of augmented individualism and instead argues for the existence of a plural subject. Her account thus moves away from a game-theory framework in which coordination arises through a "me

¹²³ Gilbert does this both in her model for joint commitments as well as her model for how political societies are created from smaller social groups. See: Gilbert, *A Theory of Political Obligation*, 93-124.

watching you watching me” dynamic with a common knowledge success condition thrown in, instead arguing that joint commitments are not reducible to individual intentions.¹²⁴ Thus, her idealizations track the creation and normative structure of a plural subject. For this reason, Gilbert’s model for shared intentions is not merely about understanding reality but also explicitly about managing it.¹²⁵ As a result, Gilbert also starts with a second-order idealization, which is that her model for joint commitments is universal in scope.¹²⁶ The idealization of universality goes further than agent-neutrality, because the model purports to capture the basic structure of joint commitments and, by extension, manner in which all people (regardless of who they are or broader context in which they exist) build their social and political world. Gilbert’s second-order idealization thus impacts what kind of first-order idealizations she employs. Here I focus on two significant first-order idealizations: (1) idealization of the average, and (2) ideals as social reality management.

One of the features of Gilbert’s argument is the claim that a joint commitment arises when two or more people express a personal readiness to enter into the commitment. The question of what constitutes personal readiness involves an idealization of human functioning. A display of personal readiness need not be a verbal statement in particular, but it must be an expression of readiness that others can interpret with a certain amount of ease in most cases.¹²⁷ Gilbert motivates this claim by idealizing average human functioning, or, to use her words,

¹²⁴ Gilbert, *Joint Commitments*, 4.

¹²⁵ I argue in Chapter III that Gilbert’s model is not unique as a form of reality management but, in contrast with Bratman’s descriptive project, it explicitly includes normativity from the start.

¹²⁶ In contrast, Bratman’s model offers a set of sufficient conditions for small-scale shared agency and does not claim that extensionality necessarily follows, leaving room for other models. Gilbert’s model thus operates at the exclusion of other possible models.

¹²⁷ If this were not true for Gilbert, then it would be hard to motivate a broader claim that joint commitments are ubiquitous and definitive of our social and political life.

‘normal’ human functioning. For Gilbert, the display of personal readiness requires that agents are ‘normal’ human beings with similar perceptual and conceptual capacities such that they can perceive each other’s expression of readiness.¹²⁸ This idealization is, in the first case, an idealization by approximation, because it purports to track how human capacities function in a statistically average way. In the case of personal readiness, the claim is that the differences between average functioning humans are irrelevant (or sufficiently negligible) when those humans interpret each other’s displays of personal readiness. However, this idealization of average functioning introduces a different kind of idealization, namely that average humans have a sufficiently shared background of hermeneutical resources, such that the differences between them can be framed as irrelevant or negligible. These idealizations are thus used to argue for the ubiquitous nature of joint commitments by showing that displays of personal readiness can be generally interpreted with ease – thus forming the basis for our entire social and political worlds.

In addition, Gilbert employs another first-order idealization in order to claim that joint commitments are necessarily normative. It is clearest here that Gilbert’s model is, in some sense, more about the management of reality than a description of it. If joint commitments are normatively binding on all participating members, then her model is action-guiding. It offers a universal model for how shared intentions ought to work and thus turns the idealization about average human functioning into an ideal for human functioning. For Gilbert, the explanatory power of her account is its ability to capture everything from small-scale shared agency to large-scale political society. Since joint commitments are also normative, this model offers a way to manage social reality by explaining how we ought to manage ourselves both by way of joint

¹²⁸ Gilbert has defined the normalcy of agents in terms of capacities for reasoning and the presence and function of perceptual organs, See: Gilbert, *A Theory of Political Obligation*, 174n64-65; Gilbert, *On Social Facts*, 188: “In particular they have normal perceptual organs functioning normally, and they have normal reasoning capacities.”

commitments and in the context of particular joint commitments. In this case Gilbert's idealizations, if viable, offer a model that strikes a valuable balance between theoretical ease and computational accuracy, because her basic model can capture the extraordinary range of complexity of our social and political life.

iii. Kutz: Rational Choice Theory

Finally, Kutz's model for participatory intentions is descriptive, insofar as it explains how participatory intentions work, and normative, insofar as his account of responsibility in cases of collective action relies on the descriptive model. Similar to Gilbert, Kutz offers a universal model that begins from basic components in order to capture with theoretical ease the complexity of collective actions and the responsibility of individuals who participate in them. In this way, Kutz's second-order idealization is similar to that of Gilbert, going beyond the kind of agent-neutrality found in Bratman's idealization of the pre-social individual to claim about how all forms of collective action work and offering a universal normative structure for the participants. However, Kutz's most significant first-order idealization is one that Gilbert explicitly rejects, namely that participatory intentions can be explained in terms of a pro-social game theory. This idealization requires Kutz to draw on rational choice theory (albeit modified to avoid a commitment to a metaphysical and psychological solipsism), which comes with its own set of idealizations.

Rational choice theory, broadly constructed, is the claim that we can understand human behavior (and by extension collective outcomes) through a particular framework for what it means to behave rationally. In this case, rationality is defined by the way in which an agent orders their preferences (based on their beliefs about the world) as well as develops and carries out goals through their actions. Traditional rational choice theory includes the idealization that

individuals develop their preferences, make decisions, and carry out goals in self-interested ways, independent of any connection to that of others. Game theory, by extension, aims to model how these self-interested agents engage in decision-making in relation to others, specifically in cases of conflict or cooperation. Kutz's modified game theory argues that the idealization of independent decision-makers is bad insofar as it cannot sufficiently explain collective forms of rationality. Thus, for Kutz, collective rationality requires that individual agents understand their choices to be interdependent, such that they can develop a shared preference through deliberation, and this shared preference is thus taken up when individuals develop a participatory intention.¹²⁹ However, Kutz's account of participatory intentions still retains an idealization about rationality found in both rational choice theory and game theory, which includes the ability to easily assess others' preferences and the ability to predicate the behavior of others.

Kutz's idealizations in this case connect to his broader methodological commitment to functionalism concerning mental states, particularly intentional states.¹³⁰ On this model, by idealizing the role of rationality in human action, we can, for the most part, successfully "attribute content to individuals' intentions by interpreting their planning and action in terms of reasons that explain and rationalize that deliberate behavior."¹³¹ While rational choice theory traditionally includes an idealization about first-person epistemological privilege in the case of knowing one's own preferences and goals, Kutz's functionalism rejects this idealization while

¹²⁹ There is a similarity here in both Kutz's and Bratman's reliance on rational choice theory in the case of shared deliberative processes – however, Kutz's claim relies on a phenomenological claim that rational choice theory cannot account for collective rationality, whereas Bratman argues that social rationality is an augmented form of individual rationality. For this reason, they both rely on idealizations implicit in rational choice theory, but modify those idealizations differently.

¹³⁰ Kutz, *Complicity*, 72-74. Kutz's functionalism is akin to that of Bratman's functionalism, which underlies his account of individual planning agency (73n17). It is worth noting that functionalism is a form of idealization because it relies on folk psychology. See: Appiah, *As If*, 51-53.

¹³¹ Kutz, *Complicity*, 74.

maintaining an idealization about our capacity to interpret the behavior of others successfully by applying a general theory of human rationality to their behavior. In this way, Kutz's useful untruth is that humans generally act in accordance with rationality. Appiah argues that this kind of idealization is one in which paints humans as impossible "Cognitive Angels," who, in this case, always act in accordance with the demands of rationality.¹³² Appiah notes that, in reality, humans frequently act irrationally and such irrationality is not necessarily a deviation from the norm of rational behavior (and may, in fact, might be the actual norm of human behavior). However, Appiah argues that this idealization is not bad if this idealization helps us understand and manage actual human behavior. In discussing a similar case – Daniel Dennett's intentional stance¹³³ – Appiah argues that by treating other people *as if* they were rational, we can predict and understand the behavior of others in a way that is sufficiently reliable.¹³⁴ Thus, Kutz's participatory intentions rely on an idealization about how collective rationality operates insofar as it frames our capacity for collective action by acting as if we can reliably understand and predict the behavior of others in order to modify our own behavior accordingly.

Section IV. General Analysis

To conclude, I want to note the underlying similarities between the idealizations used in these three models. Again, I believe the underlying similarities are not necessarily unique to these particular accounts and are likely employed through literature on shared intentions, but I do

¹³² Appiah, *As If*, 66-69, 82-85.

¹³³ Daniel Dennett, *The Intentional Stance* (Cambridge, MA: MIT Press, 1987).

¹³⁴ It is worth noting that Kutz also appeals to the Dennett's intentional stance in order to explain how it is that we can attribute intentions to groups, but it is unclear whether Kutz applies this claim to individuals as well (see: Kutz, *Complicity*, 194-195 and 195n55), despite the claim that collective forms of agency are necessarily and causally reducible to participating individuals.

not have space for that argument here. Thus, in anticipation of my argument in Chapter III, I want to focus on three similarities: (1) an idealized epistemology (and general cognitive sphere); (2) the idealization of agent-neutrality; and (3) an idealized social ontology (i.e., oppression-excluding). I've chosen these three similarities because they are the source of much debate in social, moral, and political philosophy surrounding the use of idealizations, particularly in relation to concerns raised by oppression theorists. Given that oppression theorists take the existence of structural oppression to a pressing concern that has some bearing about how we develop models to describe and manage human social behavior, my goal analyzing the use of idealizations within models for shared intention is thus about seeing whether and to what extent these models can adequately attend to oppression theorists' concerns, assuming, of course, that idealizations are themselves necessary and methodologically valuable.¹³⁵ In this way, I argue in the following chapters that our models for shared intention can better describe and manage the reality of joint actions when they rely on idealizations that can incorporate a complex set of interrelated phenomenon, which oppression theorists refer to as structural oppression. In order to assess whether the current predominant models succeed, I focus on three areas of similarities that I take to be the most valuable in for determining their success in this case.

First, each model provides a picture of human agency that functions as a useful untruth. For example, in order to apply the models to real-world cases, we must act as if humans are reliably rational in their behavior and as if differences in human capacities do not fundamentally undermine an appeal to similar or average capacities for purposes of the models. That is, human agents are idealized in terms of our capacities and our cognition in order to show how shared intentions work independent of complicating factors that exist in actual human life. Second, each

¹³⁵ For this reason, I reject the claim that the use of idealizations is the source of the problem.

model starts from a form of agent-neutrality. Agent-neutrality refers to the idealization that features of human agency can be divided into necessary and contingent features.¹³⁶ This idealization works by removing the contingent features, such as social identities, under the assumption that these contingent features should be included later. By providing a model for shared intention that focuses only on the necessary features, these theorists can each provide a highly simplified model, which captures the basic framework for shared intentions, independent of particular and complicating factors. Finally, each theorist idealizes the social ontological background in which human agents develop and exist. In the case of Gilbert and Kutz, this idealization is valuable insofar as their goal is to offer a universal model for shared intention and joint action. The particulars of social and political worlds in which agents exist are thus removed in order to provide a basic model, which can then be applied back onto a complex, non-ideal world. Bratman similarly idealizes by removing complicating features of human life, such as deception, coercion, and inequality between and among agents. In this way, and most importantly for my argument, all three models rely on idealizations that remove the existence of structural oppression as a factor in their analyses.

¹³⁶ I take this division particularly from John Rawls (in determining what information deliberators in the original position know). The idea here is that certain features like rationality are necessary while others, like one's race or gender, or the particular mechanisms of a society, are contingent. See: Rawls, *A Theory of Justice*.

Chapter III

Shared Intentions and Unjustified Idealizations

Section I. The Purpose of a Model

Analyzing the particular idealizations within models for shared intention requires us to ask what the purpose of those models is and whether they fulfill that purpose. In *The Philosophy of 'As If'*, Hans Vaihinger argues that idealizations are useful untruths insofar as their use successfully fulfills their intended purpose.¹ The relationship between idealizations and purpose can be broken down into three separate questions. The first question about purpose concerns what the model aims to do. Vaihinger argues that our models of reality generally aim at two different but inseparable ends: (1) to understand reality; and (2) to manage reality as well as ourselves within it. The second question of purpose is a function of success conditions: do the idealizations result in a model that does, in fact, help us better understand or manage reality? This is a matter of degree and shifts over time. A set of competing theories about a particular phenomenon can offer different degrees of understanding. That is, while some theories are better than others, it does not necessarily follow that the worse theories fail to fulfil a purpose (e.g., explanatory power) but rather that they do not provide the most successful path to fulfilling that purpose.² Moreover, the most successful theory at one point in time might be later replaced by

¹ Hans Vaihinger, *The Philosophy of "As If": A System of the Theoretical, Practical, and Religious Fictions of Mankind*, trans. C.K. Ogden (New York: Harcourt, Brace, and Co., 1924).

² Of course, what constitutes 'the most successful path' or, better, 'the best possible theory at the moment' is determined by particular values. For example, the principle of parsimony is a value about simplicity, and even if one accepts the principle, it is not simplicity in and of itself that determines whether we consider a theory to be successful, but rather its ability to adequately address a range of complexity in a simplified way. Moreover, the principle of parsimony might be more or less valuable depending on the context and purpose of the theory. For instance, adherence to the principle within scientific investigation will likely look different than adherence to the principle within the social sciences. Theoretical physicist Sabine Hossenfelder argues, in a similar vein, that the value of beauty (as a value connected to parsimony) is widespread within physics but has potentially impacted that

another.³ Finally, I offer a third consideration about purpose, one borne from the concerns about methodology and methods raised by oppression theorists. If a model is successful in fulfilling its purpose, there is still a need for critical reflection about whether that purpose is ethically or ideologically concerning. This concern is connected to the question of interest: whose interests does this purpose serve? It is not sufficient to ask, “What do we want this theory to do?” and “Does the theory fulfill this purpose?” because it glosses over a more critical question: “Should we take up this purpose as a goal?”

In this chapter, I focus on the three main idealizations used throughout analytic literature on shared intentions: (1) the idealization of agent-neutrality; (2) an idealized social ontology; and (3) an idealized epistemology (and cognitive sphere). I argue that Michael Bratman’s,⁴ Margaret Gilbert’s,⁵ and Christopher Kutz’s⁶ use of these idealizations are bad for two main reasons. First, the particular idealizations in question do not produce the most successful model for shared intentions and joint action. In particular, these idealizations ignore how structural oppression (and other considerations related to power) fundamentally impacts shared agency and joint action, which results in models that cannot offer understanding or manage of this impact. Thus, the problem for these models is that they rely on the wrong kinds of idealization, resulting in models that are less successful for understanding and managing shared intentions than they

kinds of models developed for the worse. See Sabine Hossenfelder, *Lost in Math: How Beauty Leads Physics Astray* (New York: Basic Books, 2018).

³ For example, one theory might be more parsimonious than other, making it easier to apply, and thus more successful in fulfilling its intended purpose. For further discussion, see: Thomas Kuhn, *The Structure of Scientific Revolutions* (Chicago: University of Chicago Press, 1962).

⁴ Michael E. Bratman, *Shared Agency: A Planning Theory of Acting Together* (Oxford: Oxford University Press, 2013).

⁵ Margaret Gilbert, *Joint Commitment: How We Make the Social World* (Oxford: Oxford University Press, 2014).

⁶ Christopher Kutz, *Complicity: Ethics and Law for a Collective Age* (Cambridge, UK: Cambridge University Press, 2010).

otherwise could be. Second, I argue that the current idealizations used within the three prominent models of shared intention are also ideologically concerning because they result in a form of reality management that reinforces structural oppression. If the intended purpose of these models is to reinforce structural oppression by managing shared intentions and joint action, then they are arguably successful in fulfilling that purpose. If this were true, then my only goal here would be to argue that this purpose is unethical and politically concerning. However, I take that Bratman, Gilbert, and Kutz do not have this as their intended purpose, so my argument here is intended to show how their particular idealizations result in models that likely fulfill this unintended purpose.

The layout of my chapter is as follows. Section II argues that oppression theorists also rely on idealizations but in a remarkably different way than the three models under examination. By showing how oppression theorists idealize, I also argue that structural oppression fundamentally impacts shared agency and should be included in the intended purpose of the model. Part of the purpose in developing a model for how to successfully share agency should thus provide us with a better understanding of this impact and a way to better manage (i.e., actively address and mitigate or lessen) this impact, under the assumption that no form of shared agency is immune to it. My claim is not that we ought to develop a single model for shared intentions that fulfills this purpose,⁷ but rather that any model for shared intention should include this as part of their purpose. To do otherwise is to risk reinforcing the mechanisms of structural oppression. Section III argues that all three models aim at both understanding and managing reality (which includes the management of ourselves) but the particular idealizations used result in models that are significantly less successful than they could otherwise be. Finally, Section IV

⁷ See Chapter I and Chapter V for further discussion of open-ended model pluralism.

connects my methodological concern about these idealizations with the ideological concerns raised by oppression theorists.

Section II. Structural Oppression and Shared Agency

My claim thus far is that structural oppression has a necessary and fundamental impact on shared agency. In this section, I explain the relationship between structural oppression and processes of social construction as it pertains to shared agency and joint action. Given that the proposed models for shared agency and joint action aim to manage one process by which people create and maintain social reality, it would be a mistake to see these models as independent of the very processes they aim to manage. That is, the processes of social construction shape not only our classifications about agency but also the schemes of classifications themselves. My concern is particularly about the role of power within processes of social construction, how these social constructions connect to structural forms of oppression in the case of agency, and what bearing this has on how theorists develop models for shared agency and joint actions. Put simply, as both Vaihinger and Appiah argue, our idealizations about reality already reflect other idealizations.⁸ That is unavoidable. However, the problem for these models is that they each fail to appreciate this fact. By understanding the relationship between social construction and structural oppression in the case of agency, we can better see how the set of idealizations in question are not ideologically neutral and thus risk reinforcing forms of structural oppression in their application.

Much of oppression theory aims to thus start its theorizing from the non-ideal conditions of the world in order to work toward the amelioration of injustice and oppression. For example,

⁸ Kwame Anthony Appiah, *As If: Idealization and Ideals* (Cambridge, MA: Harvard University Press, 2017), 49.

in *The Epistemology of Resistance*, José Medina argues that a non-ideal theory constitutes three commitments: (1) particularism; (2) empiricism and fallibilism; and (3) meliorism.⁹ Medina argues that the commitment to particularism is about attending first to concrete realities. His analysis captures well why oppression theorists eschew idealizations in myriad forms:

Idealizations tend to be partial and distorting, obscuring the heterogeneity and complexity of actual experiences and concrete practices, which is why they do not provide an adequate standpoint for the diagnosis of social problems and injustices.¹⁰

Vaihinger would agree that idealizations are partial and, depending on the type, distorting as well. However, the binary between idealization (including abstraction) and particularism (or concrete conditions) presupposed by ideal theorists and non-ideal theorists alike assumes that the latter do not engage in their own forms of idealization.

I argue that oppression theorists necessarily rely on idealizations as well, but idealize in ways that ensure that their models of social and political reality allow us to better understand and manage structural oppression.¹¹ That is, oppression theorists rely on idealizations that tend to be highly contextualized and that idealize *from* the complexities of oppression rather than *away* from them. The appeal to concrete realities is like an appeal to start from raw data; theorists still must use idealizations in order to explain or make sense of the data. For this reason, I suggest that oppression theorists' tendency to eschew idealizations is due to the fact that most

⁹ José Medina, *The Epistemology of Resistance: Gender and Racial Oppression, Epistemic Injustice, and Resistance Imaginations* (Oxford: Oxford University Press, 2013).

¹⁰ Medina, *Epistemology of Resistance*, 11.

¹¹ Appiah, *As If*, 112-172. Appiah's analysis of the use of idealization and ideals between John Rawls and Amartya Sen highlights how oppression theorists, like Medina, who argue for a justice-enhancing framework (i.e., the idea that we start from a theory of injustice in order for the purposes of amelioration rather than a single framework for the endpoint for the ideal of justice) highlights why oppression theorists are correct in their criticism of Rawls' use of idealization and ideals, but that non-ideal accounts, like Sen's, still require both, although in a much different capacity.

idealizations within moral, political, and social philosophy are often highly decontextualized and actively exclude the existence and thus the complexities of oppression. Seeing how oppression theorists idealize social reality allows us to better highlight why the idealizations used in the three models of shared agency are inadequate.

One of the first ways in which oppression theorists use idealization is in their use of hyponyms. In the case of theorizing about oppression, hyponyms include terms, such as: ‘structural oppression,’ ‘Global South,’ or ‘Western liberal subject.’ Unlike ideal theorists, oppression theorists idealize from concrete conditions by first, in the broadest sense, starting from idealizations that refer to highly complex and interrelated systems of power, which play out across different but interconnected socio-ontological landscapes. Of course, ‘structural oppression’ is too broad to be explanatorily useful in any particular model, so one would have to further contextualize this idealization by, for instance, referring to “systemic racialized homophobia within the U.S. health care system.” This is not to say that one starts from a more general idealization, but rather the ability to refer both to particular and systemic operations of power allows theorists to idealize from high complex processes in order to make them cognitively tractable – that is, they provide enhanced insight and understanding of the complexity itself. This is true even if, taking Medina’s point, a theorist starts from a more concrete issue in order to interrogate and shift the hyponyms within the literature. For instance, an idealization like ‘Global South’ might fail to adequately capture the more particular relations between the United States and Canada as settler-colonial nations, the literal and metaphorical spaces of borders,¹² and the existence of sovereign indigenous nations within this geographical space, and thus the particular analysis in this case might call in to question both (1) the

¹² Gloria Anzaldúa, *Borderlands/La Frontera: The New Mestiza* (San Francisco, CA: Spinsters/Aunt Lute Press, 1987).

usefulness of this idealization for the particular case and/or (2) its usefulness of this idealization in decolonial literature more broadly.

Second, intersectionality as a methodology offers a different way to use contextually justified idealizations, namely, through the particular value of metaphor. Metaphors can shape how theorists conceptualize the dynamic ontological processes of highly complex systems of power. In *Intersectionality as Critical Social Theory*,¹³ Patricia Hill Collins argues that Kimberlé Crenshaw's metaphor of an intersection itself constituted a paradigm shift for how many fields of study conceptualized systems of power and how people (as members of communities and social groups) were positioned (and repositioned) within them.¹⁴ For Crenshaw, the provisional metaphor of the intersection highlights how single-axis forms of discrimination law (i.e., those focused on *either* gender *or* race) could not adequately capture the multi-axis harm of discrimination in the case of *misogynoir*¹⁵ (a term that refers to anti-Black misogyny), specifically because single-axis configurations of Blackness and womanhood were associated either with Black men and white women, such that Black women tend to lose harassment cases on either front because they are not taken to be an authoritative (or default) voice for either categorization. Crenshaw offers the metaphor of the basement that further builds on the anti-segregationist mode of thinking offered by her intersection metaphor with one that makes visible the kind of anti-subordination thinking required for theorizing through multi-intersectional and

¹³ Patricia Hill Collins, *Intersectionality as Critical Social Theory* (Durham, NC: Duke University Press, 2019), 24-50.

¹⁴ Kimberlé Crenshaw, "Demarginalizing the Intersection of Race and Sex: A Black Feminist Critique of Antidiscrimination Doctrine, Feminist Theory and Antiracist Politics," *University of Chicago Legal Forum* 140, no. 1 (1989): 149.

¹⁵ Moya Bailey, "They aren't talking about me...", *Crunk Feminist Collective* (blog), March 14, 2010), <http://www.crunkfeministcollective.com/2010/03/14/they-arent-talking-about-me/>.

unjustified social hierarchies.¹⁶ The use of metaphor within intersectional methodology offers powerful forms of idealizations that enhance theorists' abilities to think through complex, co-forming systems of power that operate between different but interconnected contexts and, by extension, a multi-axis approach to thinking through identity and identity construction within ontological oppression.¹⁷ Intersectional theorists are not alone in their use of metaphor but do uniquely take up metaphor alongside the use of certain heuristics as a valuable idealizations for capturing highly complex and dynamic processes.¹⁸ In this case, such metaphors offer viable insight in their own right, not as something to later be translated into non-metaphorical terms, but as a powerful form of idealization that allows users to better understand complex structures of power in order to more adequately address them. Again, such metaphors do not aim to exclude the complexity of non-ideal conditions (as they do an ideal methodology) but are a necessary tool by which theorists can address complex concrete, non-ideal conditions.

Third, in order to address how co-forming systems of power shape: (1) who we are, the communities and social groups in which we exist (or, better yet, move between); (2) how those shifting identities impact our level of power (or access thereof), forms of privilege (or lack thereof); (3) the meanings and possibilities for our actions; and (4) the social and political dimensions of knowledge production, legitimacy, and interpretation. In this way, talk of identity can function as its own form of idealization or fiction.¹⁹ That is, unless non-ideal oppression

¹⁶ Crenshaw, "Demarginalizing the Intersection of Race and Sex," 151-152. My analysis here also draws from: Anna Carastathis, "Basements and Intersections," *Hypatia* 28, no. 4 (Fall 2013): 698-715.

¹⁷ Collins, *Intersectionality as Critical Social Theory*, 45 -50. Collins argues that an intersectional methodology focuses on, at least, six core concepts that frame how and why theorists approach the study of oppression, which are: relationality, complexity, social inequality, social justice, power, and social context.

¹⁸ Collins, *Intersectionality as Critical Social Theory*, 24-41.

¹⁹ Here I am suggesting that identity talk itself – because identity is necessary social – relies necessarily on idealizations. If it did not, then we would lack the capacity to generalize beyond a single individual or a small,

theory requires us to never abstract or generalize in talk of identity, then these terms are used as forms of fiction.²⁰

This is to say that appeals to identity thus is a kind of fiction, but the construction of fictional categories can alter how people understand and manage themselves or others. Ian Hacking refers to this as the “looping effect.”²¹ The process of classifying humans affects those being classified. Such classifications shift how those classified see themselves and how they interpret their experiences (e.g., ‘heterosexual’), how people see each other and interpret the behavior of others (e.g., ‘autistic’), and the kinds of social groups that exist (and our experiences of group membership). The looping effect occurs when the application of these classifications results in: (1) the classified tending to act in ways that can conform more strongly to the criteria of classification and thus confirm the “correctness” of the classification; (2) adding additional content to the criteria for classification (i.e., the “discovery” that most Xs also have experienced Y); or (3) results in less direct shifts in social norms or institutions.²² For example, youth-centric beauty norms increase demand for certain cosmetic products, which then reinforces the beauty norm; demands for certain kinds of technological ease in payment, e.g., credit cards, increases

highly homogeneous community. See, e.g.: Kwame Anthony Appiah, *The Ethics of Identity* (Princeton, NJ: Princeton University Press, 2005).

²⁰ In this case, the fiction is not that these identities do not really exist (I’m not making a claim either way), but rather in order to talk about these identities in any generalized way, oppression theorists do so in a manner that recognizes that there is usually not some essential feature by which these identities are defined, that identities function differently in different contexts, and the conceptual schemas in which such identities exist are not singular nor static. That is to say, their abstract use requires theorists to engage knowingly in a fiction, even when it seems that the discussion is highly localized and particularized.

²¹ Ian Hacking, *The Social Construction of What?* (Cambridge, MA: Harvard University Press, 1999).

²² Carl Elliott, “The Looping Effects of Enhancement Technologies.” *The Journal of Bioethical Inquiry* 16 (2019): 127-131. Elliott focuses specifically on self-looping effects in cases of enhancement technology (e.g., air conditioning, neurotropic prescription medications like Ritalin), noting that how individual choices can change social norms and institutional practices. Elliott also notes this in case of competition loops. If everyone is trying to get ahead in a highly competitive job market, it increases the terms of competition. For example, one publication might be sufficient for a newly minted PhD to acquire an academic job now but, given the increase of graduate students with predoctoral book contracts, the terms of competition will likely increase in more demanding ways.

the number of cashless stores, which shuts out those who cannot afford or access this technology and puts increased pressure to conform on those who can. Appiah argues that these classifications function as idealizations because we largely proceed *as if* they are true. For example, “The category “alcoholic” is useful for helping people who have difficulty with their drinking, even if it turns out, in the end, not to be a scientifically sustainable diagnostic category.”²³ That is, these classifications are not merely intended for to help their users better understand social reality, but rather they are used to manage social reality by managing contextual taxonomies of human kinds and social groups.

Oppression theorists and other non-ideal theorists often examine the way in which certain forms of classification create hierarchical-ordered social groups without throwing out the classifications entirely. In order to critique a particular form of human classification, one still must rely on it, if only to show the material effects of its application. However, the case of social categorization is often more complicated, because those classified often come to see themselves through the identity, for better or worse. For example, the medical term ‘homosexual’ was instrumental in the creation of sexuality as an identity rather than a set of practices, which allowed those similarly classified to seek each other out, create communities, and, in turn, resist the ways in which the initial classification was used to exert social control (i.e., manage) them.²⁴ In this way, forms of social categorization are idealized insofar as theorists proceed *as if* they are real, even as they critique them, because those categorizations do, in fact, shape our social reality and have material consequences.

²³ Appiah, *As If*, 205n24.

²⁴ This is why some theorists argue, as Linda Alcoff puts well: “Social categories of identity make resistance possible but always fail to identify accurately, and thus by this very fact create the need for resistance.” See: Linda Martín Alcoff, *Visible Identities: Race, Gender, and the Self* (Oxford: Oxford University Press, 2006), 77. For an example of this kind of account, see: Judith Butler, *The Psychic Life of Power: Theories of Subjection* (Stanford, CA: Stanford University Press, 1997).

That is, talk of ‘social groups’ is itself a form of idealization. Iris Marion Young’s analysis of gender as a seriality is helpful in capturing this point:²⁵

Woman is a serial collective defined neither by any common identity nor by a common set of attributes that all the individuals in the series share, but rather names a set of structural constraints and relations to practico-inert objects that condition action and its meaning.

Practico-inert objects include: social representations that create and maintain particular gender schemas through their interaction with race, class, and age schemas, compulsive heterosexuality,²⁶ and binary sex divisions in the labor force. This is to say that there is a cultural and social milieu of the way in which gender classification (and gender as a form of classification) is understood, enacted, and enforced. That is, people experience multiple serialities as a confluence, such that the constraints and practico-inert objects are not the same and/or do not function in the same way for all individuals included in the seriality.²⁷ This confluence model of seriality thus commits to an intersectional theory of oppression and avoids thinking about one’s social position as a mere mixture between abstract forms, but rather takes up points of intersection as a necessary starting point in analyzing the mechanisms and

²⁵ Iris Marion Young, “Gender as Seriality: Thinking about Women as a Social Collective,” in *Intersecting Voices: Dilemmas of Gender, Political Philosophy, and Policy* (Princeton: Princeton University Press, 1997), 12-37. Young distinguishes between a group and a seriality. A group, even a non-voluntary social group, requires self-awareness of membership and members must be united in and through the actions they undertake together. In this way, a group is organized. For my purposes here, I take ‘social groups’ to function largely as serialities that have the potential to become an organized group.

²⁶ Adrienne Rich, “Compulsory Heterosexuality and Lesbian Existence,” *Signs* 5, no. 4, Women: Sex and Sexuality (Summer 1980): 631-600.

²⁷ It is important to emphasize that identification is not a necessary part of being in a seriality. This is why, for instance, identifying as a woman is not a requirement for being a target of misogyny or being white is not a requirement for accessing some level of white privilege. This is a function of third-party categorization, not internal self-identification. This is why terms like *whiteness* or *heteronormativity* can refer both to the behavior of social groups (i.e., white people, straight people) and things like social practices and ideologies that are taken up more broadly. This is why one’s constraints can change depending on the conceptual schema for social categories present, and given the relationship between culture and subculture, there are many cases in which a person will have to simultaneously navigate different schemas. See, e.g., María Lugones, *Pilgrimages/Peregrinajes: Theorizing Coalition Against Multiple Oppressions* (Lanham, MA: Rowman & Littlefield Publishers, 2003).

phenomenological impact of co-existent, interconnected, and often mutually reinforcing structures of oppression.²⁸ As Young writes, “No individual women’s identity, then, will escape the markings of gender, but how gender marks her life is her own.”²⁹

Because social categories function as necessarily interconnected serialities, theorists must idealize them in different ways. In *Visible Identities*, Linda Alcoff distinguishes between a *public identity* and *lived subjectivity*.³⁰ We are publicly perceived as having some social identity, which shapes how others classify and interact with us *as well as* how we see and understand ourselves. Our lived subjectivity, on the other hand, is our internal experiences of selfhood; that is, our phenomenology of agency. Because our public identity and lived subjectivity do not map perfectly onto one another, we have an *agential identity*,³¹ in which the internal and external experiences interact with one another. Because oppression theorists’ discussions of agential identity, even in their more particularized forms, cannot actually capture the true complexity of these factors and their interactions, all talk of agential identity are forms of idealization.

In this way, I do think oppression theorists have been wrong to askew idealizations, because theorizing about oppression necessarily requires appealing to various forms of idealization. However, I do think that much of the pushback against idealization has highlighted well the ways in which bad idealizations in moral, political, and social philosophy do not merely

²⁸ See Chapter IV for further discussion of intersectionality as a methodology, epistemology, and theory of oppression in the case of shared agency.

²⁹ Young, “Gender as Seriality,” 33.

³⁰ Alcoff, *Visible Identities*, 92-93. See also; Ásta, *Categories We Live By: The Construction of Sex, Gender, Race, and Other Social Categories* (Oxford: Oxford University Press, 2018).

³¹ Robin Dembroff and Catharine Saint-Croix, “‘Yep, I’m Gay’: Understanding Agential Identity,” *Ergo* 6 (2019): 571-599.

result in inferior models but are also ideologically concerning.³² If our models of social reality and its features have the potential to create, reinforce, or otherwise propagate structural forms of oppression, then there is a pressing need to critically examine the ways in which theorists idealize and for what purpose.

One primary concern for models of shared intention is that they often purport to be universally or generally applicable, relying on idealizations that ignore agential identity and the impact of structural oppression on agency. Oppression theorists tend to idealize in order to actively capture and attend to the relationship between the mechanisms of social constructionism, oppression as a structural phenomenon, and the processes in which conceptual schemas shape agential identity, conceptions of agency, as well as the ways in which we share agency (or the possibilities thereof). While the three models for shared intention tend to idealize in ways that appeal to forms of neutrality, oppression theorists do the opposite by developing idealizations that: (1) semi-isolate the relevant process of social constructionism for their analysis without appealing to a neutral, “objective” reality or using these idealizations in decontextualized ways;³³ (2) use hyponyms, metaphors, and as-if identity categories that allow them to model highly complex (and thus computationally intractable) processes for the purposes of understanding managing (e.g., resisting) those processes; and (3) allow for a self-reflection in modeling of social reality in which appeals to empirical evidence are understood in relation to processes of social construction, not as a way to determine the ‘realness’ of categories but as a way to assess the role of ideology and its material impacts.

³² See, e.g., Charles W. Mills, “‘Ideal Theory’ as Ideology,” *Hypatia* 20, no. 3 (Summer 2005): 165-84; Ásta, “Categorical Injustice,” *Journal of Social Philosophy* 50, no. 4 (Winter 2019): 392-406.

³³ It is noteworthy that Vaihinger argues that the usefulness of idealizations necessarily depends on context. Thus, idealizations that function by attempting to radically decontextualize are effectively doomed to fail.

Given that my assessment of the three models of shared intention and their idealizations requires an appeal to concepts like ‘social constructionism’ and ‘structural oppression,’ particularly the relationship between the two, the remainder of this section provides an overview of how I will be using these terms going forward. Since the term ‘social constructionism’ refers to a vast spectrum of positions, my working definition is as follows: we are in some way mistaken in appealing to an ‘objective reality’ that can be known by a neutral process of inquiry (i.e. ideal observer), which results in ‘objective knowledge.’³⁴ In this case, our concepts, conceptual schemas, and our knowledge are socially constructed such that they play a determining role in our social reality *and* are determined by our social reality. Feminist philosophers, in particular, have long argued that appeals to ‘objectivity’ in philosophy and in scientific practice often reflect androcentric bias, and that the terms of social construction in the case of gender also make it hard for those who are classified as ‘men’ to see these appeals as anything but ‘neutral’ or ‘objective.’³⁵ For example, Catharine MacKinnon argues that legally defining ‘rape’ as being beyond a ‘normal level of force’ already suggests that forcing sex is appropriate (which for MacKinnon reflects a heterosexual male point of view), and that rape is illegal because it goes beyond an acceptable level of force.³⁶ Building on MacKinnon, Haslanger argues that the problem is not that ‘rape’ does not actually exist, but rather that how people define rape:

...Can depend in crucial ways on a particular group’s responses and point of view [...]. The example also shows how claims to map “objective” reality is setting up

³⁴ I use the phrase ‘in some way’ to avoid appealing to the nature of the mistake because I leave open the question the extent to which forms of realism, objectivism, and naturalism are consistent with a broader appeal to social constructionism. For my purpose here, the concern is not which theory of social construction is correct, but rather to make a more general point about why social constructivism is of relevant concern for modeling shared intentions.

³⁵ See, e.g.: Sandra Harding, *Whose Science? Whose Knowledge?: Thinking from Women’s Lives* (Ithaca, NY: Cornell University Press, 1991).

³⁶ Catharine A. MacKinnon, *Toward a Feminist Theory of the State* (Cambridge: Harvard University Press, 1989).

classification schemes can have disturbing political consequences, for the rhetoric of objectivity can serve to mask the privileging of the dominant group's interest.³⁷

While social constructionists debate to what extent there is an objective reality and, if so, what we can know of it, the important takeaway here is that our conceptual schemas and concepts are not discovered but created,³⁸ that they subject to looping effects, which also makes it harder, if not impossible, to distinguish between brute facts and social facts (given that social facts alter some brute facts).

Importantly, processes of social construction do not operate universally, but through more globalized and localized feedback mechanisms.³⁹ This is why, as I will discuss further in Chapter IV, our epistemological framework for social reality cannot appeal to a completely shared, or “objective,” reality because to do so ignores how such mechanisms are context-dependent, such that knowledge is also context-dependent, reflecting also the ways in which agents who produce that knowledge are embedded within social reality (or very different social realities). This is also why modeling social reality requires us to critically assess the ways in which facets of social reality are not shared or not equally and jointly enacted. For example, John Searle's claim that “there can be an epistemically objective social reality that is partly constituted by an ontologically subjective set of attitudes”⁴⁰ is wrong because his model idealizes away from the context-dependency of knowledge itself. That is, there is often not an objective social reality

³⁷ Haslanger, *Resisting Reality*, 103. To be clear, Haslanger is not adopting MacKinnon's general analysis of social construction and patriarchy but rather using MacKinnon's analysis as a good example of how strong pragmatic constructions work.

³⁸ This claim is consistent with both anti-realists and realists in the social construction literature. We can create a conceptual schema, for example, that we think reflects an independent reality better than another.

³⁹ These feedback mechanisms are also global in the sense that cultures are not separate (and often reference to a ‘culture’ is reference to a broad collection of ‘subcultures’ defined by some common ground like nationality or geographical location). For example, classification systems in the case of gender and race vary globally and historically, but those variations are not independent of one another.

⁴⁰ John Searle, *Mind, Language, and Society: Philosophy in the Real World* (New York: Basic Books, 1998), 113.

to which we can all easily appeal, because in order for that to be possible, we would have to be equal participants in building the social world together, such that knowledge of this social world was in no way dependent on one's position within it. One need only to appeal to the role of power in shaping the social world to show that we are not.⁴¹

Processes of social construction thus play a key role in oppression as a structural phenomenon.⁴² To define oppression as a structural phenomenon is to say that basic structures of a society result in outcomes that are not reducible to individual agents, collective agents, and/or other large-scale institutions but rather result from "the normal processes of everyday life."⁴³ Instead, the outcomes are the result of a number of factors, many of which are not within the direct or indirect control of individuals, groups, or institutions, and include a variety of social practices (e.g., cultural norms, formal, and informal rules, et cetera).⁴⁴ The causes of oppression are thus: "embedded in unquestioned norms, habits, and symbols, in the assumptions underlying institutional rules and the collective consequences of following those rules,"⁴⁵ in conjunction with intentional behavior that reinforces those outcomes.

Structural oppression is not a single mechanism but rather an umbrella term of a number of forms of oppression that operate at various structural levels, such as racism in the form of

⁴¹ For an in-depth discussion of social power, social ontology, and collective intentionality, see: Åsa Burman, *Power and Social Ontology* (Malmö, Sweden: Bokbox, 2007). (Published under Åsa Andersson.)

⁴² While 'structural oppression' is often distinguished from 'interpersonal (or individual) oppression,' here I take 'structural oppression' to also include the interpersonal forms of oppression.

⁴³ Iris Marion Young, "Five Faces of Oppression," in *Justice and the Politics of Difference* (Princeton, NJ: Princeton University Press, 1990), 41.

⁴⁴ Here I pull both from Young's account of structural injustice in *Responsibility for Justice*: Iris Marion Young, *Responsibility for Justice* (Oxford: Oxford University Press, 2011); and her analysis of structural oppression in: Young, *Justice and the Politics of Difference*.

⁴⁵ Young, "Faces of Oppression," 41.

white supremacy and sexism in the form of patriarchy. Young argues that structural oppression exists:

[When...] social processes put large groups of persons under systemic threat of domination or deprivation of the means to develop and exercise their capacities, at the same time that these processes enable others to dominate or to have a wide range of opportunities for developing and exercising capacities available to them.⁴⁶

While Young notes that there is not a single set of criteria that captures the condition of all forms of oppression,⁴⁷ there is at least one commonality:

All oppressed people suffer some inhibition of their ability to develop and exercise their capacities and express their needs, thoughts, and feelings. In that abstract sense all oppressed people face a common condition.⁴⁸

Young uses the term *social-structural position* to refer to the one's location, level of vulnerability, and *opportunity set* within a form of structural oppression.⁴⁹ Social-structural processes 'create "channels" for the actions of individuals, guiding and constraining them in certain directions.'⁵⁰ One might think of these channels in terms of J.J. Gibson's theory of affordances: social-structural processes constrain or enable possibilities for action.⁵¹ For those in more vulnerable social-structural positions, these constraints can function like Marilyn Frye's metaphor of a birdcage; it is not each single wire that prevents flight but the set of wires

⁴⁶ Young, *Responsibility for Justice*, 52.

⁴⁷ For example, Young argues for five conditions to determine whether a group is oppressed: exploitation, marginalization, powerlessness, cultural imperialism, and violence. However, none are necessary. For Young, meeting one condition is sufficient, but most oppressed groups will meet more than one.

⁴⁸ Young, "Faces of Oppression," 41.

⁴⁹ Young, *Responsibility for Justice*, 19, 45.

⁵⁰ Young, *Responsibility for Justice*, 53.

⁵¹ J. J. Gibson, "The Theory of Affordances," in *Perceiving, Action, and Knowing: Towards an Ecological Psychology*, eds. John D. Bransford and Robert Shaw (Hoboken, NJ: John Wiley & Sons, 1977), 127-142.

together.⁵² When oppression is structural, the cage is not intentionally built, but rather the outcome of a set of social processes – some of which are, in fact, intentional. Those who do not experience such constraints often see only individual wires, failing to see how even randomly positioned wires can function as a cage. In this way, many constraints on action are not imposed by others directly, but rather the accumulation of many smaller constraints exacerbating one another in the process of accumulation.

Social-structural positionings are not static because they largely rely on conceptual schemas and forms of social categorization that are subject to change over time and often do not function the same way in all contexts. What's valuable about understanding the processes of social construction within forms of structural oppression is that there are two ways in which this relationship shapes agency. First, the looping effect of social categorization can alter how individuals develop in relation to how they are categorized and how they categorize themselves. Second, this looping effect occurs also at a group level – that is, both in the creation of new social groups by shifting the terms of social categorization and in the reinforcement of existing groups. This is why addressing the role of social-structural positioning requires us to attend not simply to individuals but more particularly to individuals as members of social groups, which are hierarchically positioned relative to one another and do not operate as discrete categories.

Additionally, on this structural model, oppressed social groups need not have a correlating oppressor group, because oppression functions as the cumulative outcomes of individual practices that create and maintain the structural injustice.⁵³ However, to say that social

⁵² Marilyn Frye, *The Politics of Reality: Essays in Feminist Theory* (Freedom, California: Crossing Press, 1983), 7.

⁵³ That said, the line between privileged group and oppressor group is blurry. For instance, if there is privileged group that does not constitute an oppressor group, then it seems that the only reason it is not an oppressor group is because those in the privileged group are actively trying to undermine the hierarchy from which they benefit. Otherwise, it seems like most privileged groups are also very likely to function as oppressor groups, even if members of the group are not intentionally or consciously trying to maintain the oppression.

groups are positioned hierarchically in relation to one another is to claim that: “for every oppressed group there is a group that is *privileged* in relation to that group.”⁵⁴ In this way, the social construction of social categories like race, gender, and sexuality (and conceptual schemas in which items like ‘race,’ ‘gender’ exist) impact those at different points of confluence (i.e., social-structural positionings) differently in terms of constraints, experiences, and knowledge and ignorance creation. Moreover, forms of structural oppression reinforce the categorization in ways that often make them appear natural, rather than social, even to those who do not benefit from them.

In analyzing models of shared intention, the relationship between social construction, knowledge, and structural oppression highlights key concerns. First, if processes of social construction impact how agency is understood, particularly in relation to social categories, there is an internal and external impact. The internal impact is how one develops as an agent, one’s phenomenological sense of agency, one’s self of sense in relation to others (particularly in terms of how forms of social categorization often prescribe the possibilities of those relationships, even if not deterministically), one’s agential capacities, one’s options for action (including the meanings those actions take on within a given context), and the ease by which one can both perceive and exercise the available options. The external impact includes how our conceptual schemas shape our understanding of agency (i.e., what capacities or features are regarded as central), who counts as a possible participant in cases of shared agency, what counts as a joint action (and the structure that underlies such actions), the meaning of individual and joint actions (and, by extension, the terms of their success), and how social-structural positionings within a given context impact how differently situated agents come to share agency.

⁵⁴ Young, “Five Faces of Oppression,” 42.

Moreover, differences in conceptual schemas can result in disagreements about reality that create problems for how the goal of a shared intention is understood. There is a necessary connection between social-structural positioning, knowledge related to the mechanisms and effects of oppression, and the extent to which hermeneutical resources are shared – all of which impact the successful development of shared intentions and joint actions.⁵⁵ For this reason, when models for shared intention rely on idealizations that cannot account for related phenomena that play constitutive (not just causal) role in meaning, development, and expression of agency (and the particular capacities associated with it), then we ought to be critical about whether and the extent to which such models allow their users to understand what shared intentions are and their conditions for success. If these models manage shared intentions by defining what they are in terms of narrow success conditions, then we ought to be concerned, because this ignores how the meaning of ‘joint action’ and ‘shared agency’ is already partially formed by our conceptual schemas.

Section III. Three Models and Their Idealizations

Oppression theorists often argue that if a theory aims to effectively capture or account for structural oppression, then the use of idealizations is incompatible with this goal. That is, if a model of shared intention were to sufficiently account for the impact of structural oppression, then it could not rely on idealizations related to human agency or social ontology. However, following from Vaihinger’s claim that idealizations are both methodologically necessary and valuable, then the oppression theorists’ concern about idealizations in general can be reformulated into a concern about which idealizations better accommodate the existence and

⁵⁵ This argument appears in Section III of this chapter but appears in a more substantive form in Chapter IV.

impact of structural oppression. My goal here is thus to show that that idealizations used by the three predominate models of shared intention cannot adequately accommodate the impact of structural oppression on shared agency, but also that accounting for the concerns raised by oppression theorists requires us to use different idealizations.

The models of shared intention offered by Bratman, Gilbert, and Kutz all contain a descriptive aim: to understand what shared intentions are and provide conditions for their success. However, each model also offers a way to manage reality and ourselves. That is, in addition to a descriptive account, each model offers a normative framework that aims to guide how we share intentions and what kinds of agents we must be in order to successfully participate in them. By analyzing the idealizations used by these three models, I show how each takes on these two aims and why the current idealizations do not result in models that sufficiently fulfill these aims because they cannot account for a phenomenon that fundamentally shapes shared agency: structural oppression. Again, I focus on three similarities between the idealizations used in each model: (1) an idealized social ontology; (2) the idealization of agent-neutrality; and (3) an idealized epistemology.

i. Idealized Social Ontology

Following Section II, I start with the use of idealizations related to social ontology. In “Ideal Theory as Ideology,” Charles Mills objects broadly to the use of an idealized social ontology on the basis that such idealizations will offer little to nothing about “actual historic oppression and its legacy in the present, or current ongoing oppression, though these may be gestured at in a vague or promissory way (as something to be dealt with later).”⁵⁶ Assuming we

⁵⁶ Mills, “Ideal Theory as Ideology,” 168. It is noteworthy that Mills follows Onora O’Neill’s distinction between idealization and abstraction (1987). Under Vaihinger and Appiah’s accounts, abstraction is a form of idealization. My argument here thus still tracks Mills’ concern with a certain set of idealizations while maintaining that there are

cannot reject the use of idealizations in the case of social ontology, the question then is whether oppression-excluding idealizations are useful untruths for developing an adequate model of shared intentions. Following Vaihinger, Appiah notes that there are two reasons why idealizations are justified in excluding complicating factors: (1) when those factors are relatively inconsequential (i.e., they do not play a fundamental role); and (2) when we have reason to believe that exclusionary idealizations “succeed in corresponding to reality” such that we can, in fact, better understand and manage the phenomenon in question.⁵⁷

Following Young’s analysis, structural oppression plays a fundamental role in shaping agency, and thus models of shared intentions are only justified in using an oppression-excluding, idealized social ontology if it successfully corresponds to the reality of agency within structural oppression *and* that complicating factors about the reality of oppression can be adequately reintroduced at later stages of the theory or in the application of the model. In the remaining part of this section, using Bratman’s idealizations as an example, I argue that we are not justified in using oppression-excluding idealizations in this case because they result inferior descriptive models for understanding what shared intentions are, the conditions for their success, and their relationship to joint action.

Bratman’s model for small-scale shared agency relies on oppression-excluding idealizations by starting from cases that exclude asymmetrical power relations or hierarchies between those sharing an intention. While Bratman is merely intending to exclude justified hierarchical positioning between agents within a particular action (i.e., no agent is an official leader nor has more power than the other), his idealizations also exclude unjustified forms of

indeed other idealizations that function in accordance with Mills’ argument concerning the useful and necessary role of abstraction.

⁵⁷ Appiah, *As If*, 27.

social hierarchy and authoritative power that are definitive of oppression. This leaves in him an odd position regarding the relationship between coercion and shared intentions.

On Bratman's account, two or more people cannot successfully share an intention to paint a house together if one person has coerced the other into doing so.⁵⁸ Shared intentions must reflexively interlock.⁵⁹ This means that each agent must understand the other as an *intentional co-participant* in planning to carry out the action,⁶⁰ and the content of each's intention must refer to the presence of the other. The problem here is that structural oppression functions by way of systemic constraints, which means that coercion can occur at both interpersonal and structural levels. Within structural oppression, there are a number of forms of restrictions on people's behavior that affect the range of a person's available choices for action, how they act, and with whom they act. In Shay Welch's analysis of how women plan and act within a culture of gendered violence, she provides a helpful taxonomy of the kinds of restrictions placed on individual behaviors in the case of coercion.⁶¹ For Welch, there are individual restrictions and systemic forms of coercion. Individual restrictions appear when there is a direct threat toward the person acting or indirect threat toward those they love. This includes both threats of physical harm and social punishment. Such threats can be explicit or implied and need not be fully formed in order to be coercive. Coercion is taken to be the strongest form of restriction compared to other forms, i.e., necessity, duress, and compulsion, because the person acts voluntarily but is ultimately unfree in their action. When acting under coercion, a person is acting solely to avoid

⁵⁸ Bratman, *Shared Agency*, 37-38, 101-102.

⁵⁹ Bratman, *Shared Agency*, 48-52.

⁶⁰ Bratman, *Shared Agency*, 48-50.

⁶¹ Shay Welch, "Forms of Restrictions and Systemic Patriarchy," in *Existential Eroticism: A Feminist Approach to Women's Oppression Perpetuating Choices* (Lanham, Maryland: Lexington Books, 2015), 55-71.

the threat, which drastically alters how they deliberate about any plans for action while the threat is present. When Bratman argues that a shared intention must be without the presence of coercion, he is concerned with cases of interpersonal coercion in which the threat is direct and fully formed. At the interpersonal level, he fails to account for indirect threats or those that are not fully formed, which means he treats the presence of coercion as a binary, not as a matter of degree. Moreover, Bratman ignores how a form of structural oppression can be coercive even if there is no clear interpersonal coercion. Welch's account of patriarchy as a form of systemic coercion is helpful for understanding why this creates a problem for Bratman's account.

Forms of structural oppression, like patriarchy and white supremacy, operate as a form of coercion due to their mechanisms of operation. Those who are systemically targeted by the mechanisms of oppression do not experience this as episodic and domain-specific, but rather experience barriers constantly throughout their lives and in all domains.⁶² Moreover, because the norms and practices of our society are bound up in patriarchal oppression, systemic restrictions need not be agential; that is, systemic coercion is possible even if there is "no identifiable perpetrator inflicting the restriction on the oppressed person."⁶³ Moreover, the terms of social categorization can also function as a kind of agential restriction in themselves, which agents might accept as a natural fact about their agency (i.e., members of Group X, or Group X itself, are not compatible of action Y). In this way, Welch captures how systemic patriarchal coercion shapes how women deliberate about their choices, what choices they have, and how they act, often forcing them to engage in actions that they would not have otherwise.

⁶² Welch, *Existential Eroticism*, 62.

⁶³ Welch, *Existential Eroticism*, 63.

Consider the case of rape culture as a form of sexual terrorism within patriarchal oppression. In a society like the United States where sexual assault and violence against women is widespread, women are tasked with avoiding sexual violence. When ‘women’ are structurally oppressed, they are placed in a more vulnerable position in relation to men (at least men of the same race),⁶⁴ and thus the omnipresent threat of violence can function coercively on how women act in relation to men, even if those individual men are not acting coercively. This is why women will often place their car keys between their fingers to make a weapon while walking to their car alone at night, even if they do not encounter another person, and even if it is statistically less likely that one’s attacker is a stranger. The threat of sexual violence is implied and omnipresent throughout women’s lives, and women alter their behavior in ways they otherwise wouldn’t as a result. As Welch argues, “The threat alone, bolstered by the historical and present evidence of those inchoate dangers becoming painful reality for so many of one’s group, suffices to motivate individuals to act as the norms dictate.”⁶⁵

Bratman’s oppression-excluding idealizations thus fail to attend to the presence of systemic coercion within shared intentions, particularly in cases where one person systemically benefits from the other person’s oppression. If the presence of coercion undermines the possibility of sharing an intention, then the practical application of Bratman’s model suggests

⁶⁴ I note this because vulnerability *qua gender* alone does not capture how that vulnerability is not the same for all women and that there are ways in which some men might actually be vulnerable *qua gender* compared to white women in particular. For example, within the U.S., white women are not necessarily more vulnerable *qua gender* than men of color, particularly Black men. One might argue that Black men are more vulnerable *qua gender* than white women in many contexts because of how whiteness and anti-Black racialization intersects with gender schemas. This also affects how individuals respond to systemic coercion in terms of which men are seen as a potential threat and why (and how being perceived as a threat, in some cases, is actually what makes one vulnerable). This is to say, generally, that claims about vulnerability cannot be determined based off the isolation of one form of oppression. For a related discussion about this particular kind of case, see: Cailin O’Connor, Liam K. Bright, and Justin P. Bruner, “The Emergence of Intersectional Disadvantage,” *Social Epistemology* 33, no. 1 (2019): 23-41. For a broader discussion of multi-axis analysis, see Chapter IV.

⁶⁵ Welch, *Existential Eroticism*, 64.

that shared intentions only occur when there is no structurally produced power imbalance between agents (which I take to be impossible, or incredibly rare) or, more likely, that they do not exist at all under conditions of structural oppression. Neither I nor Bratman want to argue that either of these options is the case.⁶⁶

Thus, Bratman's model does not provide a sufficient understanding of how shared intentions work and what conditions determine their success, because his oppression-excluding idealizations in the case of social ontology are not, in fact, useful untruths. The problem is that these idealizations provide us with a model of shared intentions that helps us understand a reality in which the stipulated counter-factual are, in fact, true. Both Vaihinger and Appiah reject this kind of idealization, because there is no pragmatic value in being able to understand and manage an impossible reality.⁶⁷

ii. Idealization of Agent-Neutrality

Charles Mills also argues that oppression-excluding idealizations in the case of social ontology impact the idealizations about the metaphysics of agency, which I refer to as the idealization of agent-neutrality. An additional reason Mills objects to this kind of idealized social ontology is that it posits a contestable picture of human agency insofar as it excludes how agency is shaped by social ontology:

An idealized social ontology... will typically assume the abstract and undifferentiated equal atomic individuals of classical liberalism. Thus it will abstract *away* from relations of structural domination, exploitation, coercion, and oppression, which in reality, of course, will profoundly shape the ontology of those same individuals, locating them in superior and inferior positions in social hierarchies of various kinds.⁶⁸

⁶⁶ Bratman, *Shared Agency*, 3-39. Bratman conceives of shared intentions as quite commonplace.

⁶⁷ Appiah, *As If*, 16.

⁶⁸ Mills, "'Ideal Theory' as Ideal Theory," 168.

The idealization of agent-neutrality thus starts from the untruth that agents are pre-social and that pre-social individuals are fundamentally the same. While an oppression-excluding, idealized social ontology ignores the impact that structural oppression has on agency by way of constraints, it also ignores how agency develops under conditions of structural oppression (i.e., in relation to one's social-structural positioning), not apart from them. Lisa Schwartzman captures this problem further:

... Because oppression is a group-based phenomenon, social forces must construct, shape, and mold people into different categories, or groups. In fact, these groups are often so well-defined that they appear natural, which makes people less likely to question or challenge the system of oppression they serve.⁶⁹

Because categories of race and gender are not natural categories, the formation of social categories can either intersect with oppression or be used as the basis for oppression. The problem is that in order for a form of oppression to justify itself, it must posit these social categories as natural and, by extension, as naturally hierarchical. It is in this way that a form of oppression creates hierarchically positioned social categories and shapes people in relation to them while simultaneously presenting these categories as natural and unchangeable.

The idealization of agent-neutrality presents a two-fold problem for models that rely on it. First, it assumes that the role of social categories under structural oppression play an insignificant role in developing a descriptively useful model such that we are justified in excluding them (this is especially for Gilbert's and Kutz's models that aim to be universally descriptive). I take this assumption to be incorrect.⁷⁰ However, the idealization of agent-

⁶⁹ Lisa Schwartzman, *Challenging Liberalism: Feminism as Political Critique* (University Park, Pennsylvania: Penn State Press, 2006), 168.

⁷⁰ See, e.g.: Alcoff, *Visible Identities: Race, Gender, and the Self* ((2006); Sandra Lee Bartky, *Femininity and Domination: Studies in the Phenomenology of Oppression* (New York: Routledge, 1990).

neutrality runs into a more pressing problem. In order to abstract away in this manner, we require an adequate understanding of what we are abstracting away from, otherwise we cannot adequately reintroduce those complexities while ensuring the model is useful by Vaihinger's standards. If social categories were natural and unchangeable, then the idealization of agent-neutrality would likely result in a model that could be later extended to incorporate the further complexity presupposed by these categories.⁷¹ However, social categories are not natural nor static, but because the mechanisms of oppression present them as such, there is reason to be concerned with our present ability to abstract away from them because the mechanisms of oppression tend to obscure how these categories form and operate.

Moreover, even without obscurity, the dynamicity of social categories raises a question about whether the idealization of agent-neutrality is a useful untruth. This idealization requires us to act as if there is a universal model for human agency that provides an adequate basis for understanding and managing our agency, independent of the constructed differences between agents. However, if those differences are necessary for understanding human agency, then the idealization requiring the lack of difference will likely result in an inferior model. If the goal of the model is to manage human agency, then we ought to be concerned about what that management means, given the relationship between social categories and structural oppression.⁷²

⁷¹ Of course, this is likely to be true only if theorists are also justified in using an oppression-excluding, idealized social ontology.

⁷² Appiah, *As If*, 182n29. Appiah notes that: Rae Langton raised the objection that "this account might justify sexist or racist beliefs, given that such beliefs help men control women, or one race control another." Appiah argues that this cannot happen on Vaihinger's account, because many of the elements of sexist or racist beliefs are not helpful for controlling the world, but that widespread acceptance of the beliefs can maintain structural oppression (which, for Vaihinger, are not the same thing). However, I think Appiah is mistaken in this case. Langton's objection seems correct: for example, hierarchical racial categorization in South Africa did function as useful means of control for the establishment and maintenance of apartheid. Appiah fails to appreciate the creation of social categories can result in seemingly natural categories in which beliefs about those categories can, in fact, help particular individuals manage the world. For instance, when heterosexual women and men internalize related gender roles and thus are more likely to act in accordance with them, those beliefs about the 'realness' of those gender roles might help both groups more successfully manage their dating practices (even if third parties have good reason to critique those roles).

This is particularly concerning given that social categories themselves are products of ways in which people have already tried to understand and manage reality – for better and for worse.

In the case of structural oppression, social categories have been used to create different groups that are hierarchically ordered in order to manage people in unjustified ways. This hierarchical ordering also creates a dynamic in which certain social categories are treated as neutral insofar as they are seen as the default by which human difference is marked. The idealization of agent-neutrality risks uncritically adopting a supposedly neutral picture of agency which, in fact, is a reflection of particular social categories. Moreover, those particular social categories are often those whose members largely benefit from the processes of structural oppression. For example, social categories such as ‘white’ or ‘heterosexual’ are frequently seen as the absence of a social category rather than as manufactured social categories in their own right. Because the idealization of agent-neutrality is not actually neutral, the attempt to reintroduce complicating factors like structural oppression and social categories is often the reintroduction of those social categories that are seen as different. For this reason, the idealization of agent neutrality only works if social categories were natural categories; otherwise reintroducing the impact such categories can have on shared agency treats the non-neutral ‘neutral agent’ as natural (and thus necessary) and all others as social (and thus contingent).⁷³ In this way, the idealization of agent-neutrality results in a model that uncritically reintroduces

or practices). This is why it is not merely enough that a model helps us more effectively manage reality or ourselves; we also need to address ethical and political concerns related to management.

⁷³ Here I draw explicitly from the way in which John Rawls idealizes human agents as deliberators in the original position: John Rawls, *A Theory of Justice, Revised Edition*. (Cambridge, MA: Belknap Press of Harvard University Press, 1999). He effectively idealizes agency by breaking the features of agency down into two categories: necessary or contingent. Necessary features – like rational self-interest and a sense of justice – are supposed to ensure that the outcome of the process does not contain forms of bias related to contingent features, such as social categories or social structures. However, one concern with this division is that it already reflects a kind of particular bias about how to make the division in the first place – that is, the division itself is not neutral.

human difference, leaving users of the model unable to account for how difference itself is socially constructed and unable to critique the model's presumptions about 'neutrality' and 'difference.' We need to be able to account for how social categories impact agency (and shared agency) and contest the terms of those social categories at the same time. If these agent-neutral models are useful in Vaihinger's sense, then they are useful because they build on and reinforce hierarchically ordered social categories.

Developing a model of shared intentions that relies on both oppression-excluding, idealized social ontology and the idealization of agent-neutrality will not be useful for capturing how shared intention arises between different agents in which 'difference' is a function of social categories, especially because of how 'neutrality' functions. Consider the following analogy: imagine using a social scientific methodology to study marital success that starts from the assumption that marriage is a universal human practice and the differences between this practice over time or across cultures are irrelevant to understanding how two or more people have a successful marriage. In this case, one would rightly point out that marriage is not a universal practice, although it is a common one, and it is only by understanding how this practice operates in particular contexts that one can determine whether a marriage is successful. That is, the measure of success is determined by the particular conditions because what constitutes a marriage, and thus a successful one, is determined by a particular context. Moreover, the model for a successful marriage in one cultural context cannot be used to determine the conditions for success in others.

Similarly, we ought to be concerned with universal models for shared intention – such as those of Gilbert and Kutz – because they assume that is one way to engage in shared agency. In this way, both Gilbert's and Kutz's models also manage reality (and ourselves). By arguing that

their models for shared intention are universal, they also are making a claim about what form all shared intentions should take. That is, if there is more than one way of sharing an intention, these models actively exclude those possibilities.

This exclusion plays out in two ways. First, these models will exclude certain cases that seem to resemble examples of a successfully shared intention and joint action but fail to match their model. In this case, the model will label them as unsuccessful or simply as not shared intentions, and, because the models are universal, they will also exclude the potential for other models of shared intention that might be able to provide an understanding of these excluded cases and the conditions for their success. Second, the exclusion might fail to capture cases of shared intention that do not resemble the model at all. This kind of exclusion produces a bad (i.e., not merely inferior) model because it decreases our understanding of shared intentions by using idealizations that exclude the possibility of different kinds of shared intention prematurely (i.e., before we have good reason to exclude certain cases). That is, if there is a possibility of different kinds of shared intention, this model prematurely excludes this possibility, leaving us without the ability to investigate further. Moreover, the use of a universal model imposes a structure for what it means to share an intention, which like the creation of social categories, can impact how people try to share intentions. By excluding the possibility of different kinds of shared intention, these models risk shifting how people actually try to share intentions and using these cases of imposed conformity to the model as further evidence for the universality of the model itself.⁷⁴ Given that an oppression-excluding, idealized social ontology and the idealization

⁷⁴ This follows from Ian Hacking's account of the 'looping effect' but data ethicists also refer to this as a 'toxic loop.' Toxic loops occur when contextually unjustified idealizations are used and when the application of the model does not provide feedback for the usefulness of the model and/or the model is deemed a success by similarly reinforcing data points that already reflect parochial values or conceptual schemas used in the model. See: Cathy O'Neill, *Weapons of Math Destruction: How Big Data Increase Inequality and Threatens Democracy* (New York: Crown Publishing, 2016), 32-49.

of agent-neutrality already risk reinforcing structural oppression, the concern extends to how we try manage, not merely understand, shared intentions because these models can alter the potential possibilities for how people successfully share intentions in ways that, again, reinforce forms of oppression.⁷⁵

iii. Idealized Epistemology

One concern that many oppression theorists have in the use of idealizations are those related to human agency, particularly in relation to agential capacities and social cognition. Onora O’Neill argues that idealizations rely on counterfactual stipulations, such as *homo economicus*, introducing a false predicate into the theory, making in inapplicable to actual humans.⁷⁶ Following O’Neill, Charles Mills argues that idealizations about human agency ignore or gloss over the non-ideal realities of living under structural oppression, noting that those who approximate the ideal are likely those not subject (or less subject to) oppression:

The human agents as visualized in the theory will also often have completely unrealistic capacities attributed to them—unrealistic even for the privileged minority, let alone those subordinated in different ways, who would not have had an equal opportunity for their natural capacities to develop, and who would in fact typically be disabled in crucial respects...

Separate from, and in addition to, the idealization of human capacities, what could be termed an idealized cognitive sphere will also be presupposed. In other words, as a corollary of the general ignoring of oppression, the consequences of oppression for the social cognition of these agents, both the advantaged and the disadvantaged, will typically not be recognized, let alone theorized. A general

⁷⁵ For example, imagine a case of a successfully shared intention that is wrongly (and perhaps perniciously) labeled as ‘mob behavior’ rather than a case of genuinely shared agency. This not only fails to appreciate the success and understand the conditions of that success, but also enforces a claim about what kinds of behavior count as shared agency and thus encourages agents to act in accordance with the success conditions of the model, instead of in accordance with a form of shared agency that has been successful (but is excluded and perniciously mislabeled by the model). This is of particular concern when trying to understand how shared agency functions successfully within anti-oppression movements, especially if a supposedly universal model is to discredit or discount forms of shared agency developed by those in more marginalized social-structural positionings.

⁷⁶ Onora O’Neill, “Abstraction, Idealization, and Ideology in Ethics,” in *Moral Philosophy and Contemporary Problems*, ed. J.D.G. Evans (Cambridge, UK: Cambridge University Press, 1987), 55-65.

social transparency will be presumed, with cognitive obstacles minimized as limited to biases of self-interest or the intrinsic difficulties of understanding the world, and little or no attention paid to the distinctive role of hegemonic ideologies and group-specific experience in distorting our perceptions and conceptions of the social order.⁷⁷

The three models for shared intention idealize human capacities and cognition in a similar vein, particularly in the construction of their success conditions. In order for two or more people to successfully share an intention, the participating people must have mutual knowledge or understanding of intention, which rests on the ability to correctly interpret and predicate the behavior of another. For this reason, the viability of these idealizations turns on whether they are inapplicable to counterfactual stipulations, as O'Neill argues,⁷⁸ or whether they provide a sufficient basis for interpretation and predication. My focus is thus on how Bratman, Gilbert, and Kutz idealize human rationality and social cognition (particularly agents' social epistemic capacities) and to what extent these idealizations are valuable for this end.

Appiah refers to idealizations concerning human capacity and cognition as a stipulation about Cognitive Angels.⁷⁹ 'Cognitive Angels' refers to a false picture of human agents who live up to the demands of ideal rationality and are computationally perfect such that their beliefs are true and their behavior is perfectly predictable. Appiah argues that while this stipulation is logically impossible, it does not follow that it is a useless idealization. That is, this idealization might be valuable for some models of shared intention, but not others: it depends on what the purpose of the model is.

⁷⁷ Mills, "'Ideal Theory' as Ideal Theory," 168-169.

⁷⁸ Vaihinger argues that if the counterfactuals only provide us with an understanding of a counterfactual world, then they are not useful. In this case, Onora O'Neill argues that the use of counterfactuals necessarily prevents us from applying the model to the actual world, whereas Vaihinger and Appiah maintain that counterfactuals are valuable when used correctly.

⁷⁹ Appiah, *As If*, 43.

My focus here is on success conditions in the three models that require agents to successfully interpret the mental states of others (i.e., belief, personal readiness) as well as accurately interpret and predicate the behavior of others (i.e., mutual understanding, strategic responsiveness). I start with how the Cognitive Angel idealization appears in a related context: Daniel Dennett's intentional stance.⁸⁰ In order to take the intentional stance toward another human agent, one must act as if the other person is ideally rational; that is, act as if the other person has no impediments to their functioning rationally, such that they are functioning ideally (i.e., it is a Galilean idealization). This idealization is two-fold: first, we act as if people are intentional systems; and second, we also act as if people have beliefs and desires that have bearing on their actions. That is, whether those two claims are actually true (i.e. we are intentional systems, we have beliefs and desires), it provides us with a reliable enough model to predict and interpret the behavior of others. As Appiah argues:

The strategy of predication – the strategy we use to make sense of the behavior of all the things we can usefully treat as intentional systems – is to apply an idealized model, knowing that, because it *is* idealized, it won't always get things right. To say that something sorta believes is to say that the idealized model works well enough for practical purposes, in ordinary circumstances, with that thing. If it worked perfectly, it would be just plain true that the thing believes; if it works badly enough, it's plain false. It between is a vaguely delineated world of the sorta true.⁸¹

Moreover, our as-if beliefs are a kind of bounded make-believe. To use Appiah's example of watching *Hamlet*, he argues that his sadness concerning Ophelia's death:

...Involves not an abandonment of the belief that one has died, but abandonment of one of the normal consequences of that belief, which would be (other things being equal) that I had nothing to be sad about. That's what it is to permit myself to feel *as if* someone has died.⁸²

⁸⁰ Daniel Dennett, *The Intentional Stance* (Cambridge, MA: MIT Press, 1987).

⁸¹ Appiah, *As If*, 45.

⁸² Appiah, *As If*, 108.

Idealizing human capacities and cognition operates similarly. In treating someone as an intentional system, we abandon the consequences of believing that humans are often irrational and act on false beliefs, not the belief that humans are actually irrational. In doing so, we can apply a model within a bounded context that is reliable for navigating our interactions with others.

I focus on the intentional stance, because both Kutz and Gilbert take up a version of it in their models for shared intention; that is, their accounts require similar idealizations, which require agents to act as if we can reliably predicate the belief states and behavior of others.⁸³ The success conditions for both joint commitments and participatory intentions require one to interpret and predict the behavior of others successfully. The threshold for successful interpretation and predication is relatively minimal. For Gilbert, one need only to interpret whether another has displayed personal quasi-readiness to enter into the joint commitment and ensure that these expressions are common knowledge for all involved. Kutz's threshold is weaker insofar as he replaces a common knowledge condition with the condition of mutual openness, but his condition of strategic responsiveness requires that an agent be able to reliably apply an intentional stance model to others.⁸⁴

Again, it is noteworthy that both Gilbert and Kutz aim to offer universal models – i.e., one that captures the entire range and complexity of joint actions – because that is precisely where reliance on these idealizations results in a poor model. Kutz's requirement of mutual

⁸³ Bratman's theory of planning agency does not operate off an instrumental desire-belief model for agency and thus I avoid comparing his model to Dennett here. However, in avoiding an instrumental desire-belief model, Bratman's model for shared planning agency becomes more explicitly normative – that is, there is a way in which one ought to order their desires over time in a coherent manner (and, thus, there is an implicit claim that some desires are better than others), which is why his model fulfills two aims: to understand and manage the way in which we built the social world.

⁸⁴ Kutz, *Complicity*, 76. Strategic responsiveness occurs when an agent's intention "is sensitive to their beliefs or predications about what others intend to do."

openness relies on Sperber's and Wilson's concept of *mutual manifestness*,⁸⁵ which occurs only when agents have shared cognitive background that is not the explicit content of the beliefs but can help disambiguate instances of ambiguity.⁸⁶ Similarly, Gilbert's common knowledge conditions about personal readiness require a shared conceptual background between agents as well as that agents' perceptual and reasoning capacities be functioning normally.⁸⁷ I point out here that if their models were not purportedly universal, then they might be valuable for explaining how agents who significantly share a cognitive background and who are statically normal in their perceptual and reasoning faculties come to successfully share intentions. That is, these models might be valuable for a far more limited context. However, given that these models are not limited in this way, these idealizations result in models that cannot adequately address significant differences between humans.

One reason that Dennett's model becomes less reliable is not about failures in human rationality but that successful interpretation and predication often requires agents to have a shared conceptual background and shared norms, which makes applying the intentional stance seem relatively simple and sufficiently reliable. However, the requirement of a shared background ignores substantive communicative barriers between agents between and/or within languages as well as interpretative differences in meaning of bodily gestures, which, again, are necessary components of meeting these models' success conditions. Luvell Anderson's concept

⁸⁵ Kutz, *Complicity*, 274n22. See: Dan Sperber and Dierdre Wilson, *Relevance* (Cambridge, MA: Harvard Press, 1986), 38-45.

⁸⁶ Kutz, *Complicity*, 274n22.

⁸⁷ As in Chapter II, I take this claim to be about statically normal functioning. However, if the claim about normality is not statistical but normative, then this actually creates further problems for Gilbert's idealizations in this case, which I do not have the space to detail here.

of a *hermeneutical impasse* captures this problem well.⁸⁸ Hermeneutical impasses are “instances in which agents engaged in communicative exchange are unable to achieve understanding due to a gap in shared hermeneutical resources.”⁸⁹ Such gaps can include a lack of shared language, differences in vernacular or dialectic within language, unequal skill sets of sense-making (in the case of a difficult text for example), and epistemological differences that affect a hearer’s interpretation of linguistic content. To use Anderson’s example: “‘All lives matter’ responders attribute an *exclusive* reading to ‘Black lives matter’ when an *inclusive* reading is what is intended.”⁹⁰ In this case, the failure by ‘all lives matter’ responders (who are largely white Americans) to interpret correctly is due in part to prejudices about the speakers of ‘Black lives matter’ (who are largely Black Americans), which is often linked to larger problems within social epistemology, namely epistemic injustice and ignorance.⁹¹ When purportedly universal models of shared intention require social transparency between agents in their success conditions, they become less useful (or perhaps simply not useful) for understanding how the non-idealized functioning of social cognition affects whether agents can try to share intentions, how successful they might be in these attempts, and what additional work has to occur when the agents in question do not share hermeneutical resources, for example, especially when the gap is directly related to their social-structural positioning with a system of oppression.

Moreover, even when a shared background is present to a greater extent, this kind of idealization ignores the impact of cognitive bias both in how we understand our own behavior

⁸⁸Luvell Anderson, “Hermeneutical Impasses,” *Philosophical Topics* 45, no. 2 (2017): 1-19.

⁸⁹ Anderson, “Hermeneutical Impasses,” 3.

⁹⁰ Anderson, “Hermeneutical Impasses,” 3.

⁹¹ See, e.g.: Sandra Harding, ed., *The Feminist Standpoint Theory Reader: Intellectual and Political Controversies* (New York: Routledge, 2004); Shannon Sullivan, and Nancy Tuana, eds., *Race and Epistemologies of Ignorance* (Albany, NY: SUNY Press, 2007).

and in how we interpret the intentions or the behaviors of others, and thus cannot provide us with a model that allows us to better understand and manage this impact. For example, the role of implicit bias can unconsciously impact how trustworthy we perceive others to be such that our capacity to successfully coordinate our sub-plans or determinate whether someone is expressing personal readiness is enhanced or curtailed relative to the other agent. In the case of reading another person's intended behavior, implicit bias can impact what kind of action we take the other person to be undertaking. For example, implicit bias in the case of race can result in people misinterpreting behavior of racial minorities, turning actual attempts to share agency or merely coordinate behavior into cases where one person is interpreted as engaging in anti-social behavior and thus other people wrongfully act in accordance with their misperception.

In this way, when models for shared intention require a shared conceptual, cognitive, and normative background, they also fail to appreciate how a lack of shared background affects how the action is understood and the ease by which agents can achieve common knowledge or mutual understanding.⁹² That is, they idealize epistemic capacities such that an agent's social location has no bearing on what kinds of knowledge they have, what knowledge is accessible and how accessible it is, and how agents come to develop and/or access knowledge.⁹³ Again, if the model is aiming to capture how relatively homogenous agents with a significantly shared background successfully share an intention, then idealizing epistemic capacities in this way might be valuable. However, I assume that the purpose of these models is to explain how we can successfully share intentions in both homogenous and heterogenous cases. Given that

⁹² Both Bratman and Gilbert assume that agents have common understanding with respect to the shared activity itself. If two agents do not have a shared conception of what it means to *J* – either as individuals or jointly – then this epistemic barrier cannot be worked out in sub-plans or further coordination once the joint commitment is made. Rather, we are required to ensure mutual understanding prior to attempting to share the intention in the first place.

⁹³ For in-depth analysis about non-idealized epistemological practices, see Chapter IV.

heterogenous cases might be more common, then it seems that a universal model would need to be able to capture these cases as well. The question of how we ought to idealize epistemic capacities has direct bearing on the question of how we develop a model for shared intention that can provide a better understanding of how shared intentions successfully arise given the differences produced by the mechanisms of structural oppression.

In Bratman's case, the idealized rationality operates more significantly because his model for planning agency is more normative than an instrumental belief-desire model. On his account, rational planning agency is evaluative in light terms of universal norms, which means that third parties can assess whether an agent is rational. In his critique of Bratman's account, Pierre Demeulenaere notes the relationship between norms of rationality, social norms, and the development of individual (or shared) intentions.⁹⁴ When we consider examples like going to New York together or taking a walk together, we cannot understand how such agents share an intention without understanding the social dimension that shapes and influences the intention itself. For one, our reasons are not solipsistic – our individual intentions develop in relation to social norms because social norms “delineate the legitimate sphere of preferences.”⁹⁵ The question of why we are going to New York together becomes relevant for understanding the development of our shared intention. Social norms affect our shared intentions to engage in a public activity together.⁹⁶ For example, in *Born a Crime*,⁹⁷ comedian Trevor Noah, whose father

⁹⁴ Pierre Demeulenaere, “Where Is the Social?,” in *Philosophy of the Social Sciences: Philosophical Theory and Scientific Practice*, ed. Chris Mantzavinos (Cambridge: Cambridge University Press, 2009), 60-66.

⁹⁵ Demeulenaere, “Where Is the Social?,” 62.

⁹⁶ Demeulenaere, “Where Is the Social?,” 62. Demeulenaere notes the range of possible social norms that determine shared activity in public. For example: the general norms for walking together in public, norms concerning who can walk together and how they ought to walk together, and so forth.

⁹⁷ Trevor Noah, *Born a Crime: Stories from a South African Childhood* (New York: John Murray, 2017).

was a white man from Switzerland and whose mother was a Black Xhosa woman from South Africa, details how his existence as a mixed-race child born under South African apartheid impacted his ability to walk with his mother in public. Because Noah's birth was illegal under the apartheid and because his racial categorization was different from his mother's, he and his mother could never walk side-by-side in public in order to avoid police scrutiny about his existence.

Social, cultural, and legal norms thus influence how individuals develop their intentions, although not deterministically. We might share an intention to engage in public behavior that flouts norms. Perhaps we do this because we are simply operating off of a different set of norms, or perhaps we do so with the intention of challenging or mocking those norms. What is noteworthy is that even though the influence is not deterministic, social and cultural norms also impact our assessment of agent-rationality. It is not surprising that those who intentionally flout mainstream norms are sometimes regarded as irrational and usually referred to by others as 'crazy' or categorized as mentally ill. This is true even if their behavior is consistent with other cultural or subcultural norms. To idealize rationality and its norms as purely universal and asocial results in models that cannot account for how non-universal social norms impact the development and expression of individual norms of rationality as well as how third parties determine whether another agent is rational. As Demeulenaere argues, the norms of rationality do not come prior to social norms. A particular, culturally specific conception of rationality can impact individual behavior, so it would be a mistake to extrapolate universal norms from a particular context. More to the point; "rational behavior is itself, also, a social norm that is reinforced in specific social environments and that leads an individual to conform to rational

attitudes.”⁹⁸ When individuals fail to conform, they are often regarded as irrational by others. In this way, Bratman’s idealization about rationality (and how one’s preferences order to form a rational life plan) does not merely aim to describe shared planning agency but manages it by imposing a social norm concerning rational behavior. That is, even if his success conditions are sufficient but not necessary, his model for shared planning agency necessarily builds off a model for how individual planning agency *ought* to function.

In line with Demeulenaere’s critique, Shay Welch argues that rational choice theory fails to consider how oppressed peoples develop and deliberate about their preferences particularly because it emphasizes third-party intelligibility.⁹⁹ She develops the concept of *desperate rationality* to capture how women, in particular, deliberate and act in relation to interpersonal and systemic threats of violence. Broadly, she argues that the criterion of third-party intelligibility often treats traumatized women who are dealing with ongoing interpersonal violence as behaving irrationally because those who are not subject to violence of this sort lack the relevant *knowledge of danger* – the deliberative standpoint for those experiencing such violence.¹⁰⁰ Those who lack this knowledge often regard women as acting irrationally when, in fact, they are acting rationally, using game theory strategically to ensure their survival. Her account of trauma and deliberation also focuses on the ways in which agents whose autonomy is significantly compromised or who may be operating off of adaptive preferences can still deliberate and act rationally.¹⁰¹ Standard rational theory choice thus often labels certain

⁹⁸ Demeulenaere, “Where Is the Social?,” 65.

⁹⁹ Welch, *Existential Eroticism*, 129-152.

¹⁰⁰ Welch, *Existential Eroticism*, 141-2.

¹⁰¹ In fact, Welch argues that those whose autonomy is comprised by oppressive violence have heightened rationality in many cases.

individuals as irrational, emphasizing their survival in term of luck, not rationality, and does so in virtue of its idealizations.

Section IV. Ideology and Idealizations

My argument thus far has focused on what these models take as their purpose and the extent to which they fulfill that purpose. The value of starting from Vaihinger's and Appiah's analysis of idealizations is that multiple and inconsistent models provide us with different ways of understanding and managing a phenomenon.¹⁰² That is, my analysis of Bratman's, Gilbert's, and Kutz's models has been to show why the idealizations on which their models rely do not fulfill the purpose of these models by also noting how their models might be useful for different purposes, e.g., the capacity to share intentions in homogenous populations. However, the question of whether we ought to take up these purposes still remains. Moreover, if the goal of these models is universal (as in the case of Gilbert and Kutz) or requires us to idealize away factors that cannot be reintroduced later (Bratman most significantly), then there is good reason to be concerned as to whether these models are useful for understanding a significant portion of our social reality, as they all purport to do.

Beyond understanding, all three models offer a way to manage shared intentions, which are in themselves about the management of our reality. In this way, the models are not meant to be purely descriptive because each relies on its own normative components. Gilbert's model tells us not just what a joint commitment is but how joint commitments ought to function in terms of obligation. Kutz's model is explicitly moral insofar as its purpose is to account for how participatory intentions work in order to hold people responsible for them, but the idealization of

¹⁰² Further discussion of model pluralism found in Chapters IV and V.

agent-neutrality prevents us from addressing the relationship between social categories and structural oppression as it arises within collective actions (and thus offers a model that treats mechanisms of structural oppression as irrelevant for ethical analysis). Finally, Bratman's shared planning agency is normative insofar as it appeals to norms for social rationality that frame what a joint action is and how agents ought to act within one.

Thus, the question of purpose becomes of ethical concern when our models for a phenomenon that is directly about social reality management cannot incorporate oppression – a phenomenon borne from unjustified attempts to manage our world, ourselves, and others, often by way of joint actions. If a model's purpose is to capture very low-level forms of shared agency, such as moving a table together, and does not incorporate oppression, then it seems that those models might be useful for that limited purpose.¹⁰³ However, if the model is purporting to capture a significant portion of our social, ethical, and political worlds (particularly in the way we act together in creating or maintaining them), then the purpose of these models should be able to account for how previous attempts at reality management (which have been unjust and have resulted in unjust structural mechanism) impact our current joint projects. This is what Charles Mills means when he broadly refers to idealizations as ideological:¹⁰⁴ our idealizations are not born in a vacuum, they already reflect our values (i.e., for what purpose are we idealizing?), our metaphysical assumptions (i.e., how do we idealize agency?), and our assumptions about what can be bracketed for later and what cannot. Following Vaihinger and Appiah, the problem is not the use of idealizations but rather how we determine which idealizations to use and what values or purposes those choices reflect. In this way, idealizations connect to our ideals: how ought we

¹⁰³ That said, I am generally wary of arguing that even actions like moving a table lack a need for concern about the role of oppression in such cases.

¹⁰⁴ Mills, "Ideal Theory as Ideology."

manage ourselves and how ought we manage social reality? In the case of shared agency, those questions take on an additional dimension because the question is also about who ‘we’ refers too, how this management happens by way of shared agency, who is excluded from shared agency by our models of it, and how do our social-structural positionings within structures of oppression alter our ability to successfully share agency with others. The reason why we ought to take care in how we idealize social ontology, agents, and their capacities is not simply because, as Demeulenaere argues contra Bratman that ‘the social is always already there,’¹⁰⁵ but because our social world already impacts the kinds of models we can construct and the purpose of those models. Moreover, the question of how we ought to manage shared intentions (as a form of social reality management) in light of ongoing mismanagement (or malicious management) is a pressing ethical and political question. This means that our attempts to understand how shared intentions work requires us to address how ongoing forms of reality management already impact our attempts to describe the phenomenon, how values are already reflected in our attempts at managing it, and whether we ought to take up those values.

¹⁰⁵ Bratman, *Shared Agency*, 30. Bratman pulls this quote from Demeulenaere to specifically highlight that his account does not ignore it.

Chapter IV

Modeling Shared Resistance

Section I. Whose Purpose? Whose Values?

The tradition of feminist philosophy of science emphasizes the way in which a methodology or set of methods necessarily reflects a set of values.¹ This starts in the choice of what to research.² On Hans Vaihinger's account of methodology, the way we model the world is already a function of what phenomena people take to be most pressing in terms of understanding and management.³ What the tradition of feminist philosophy of science has offered is the claim that there is no consensus about which phenomena are most pressing in terms of research. Given that humans have limited resources for research, there is already disagreement about which phenomena to model before we even arrive at the question of how to model it and whether a multiplicity of models (or open-ended model pluralism) is needed.⁴ The choice of what to

¹ For further reading, see, e.g.: Evelyn Fox Keller, *Reflections on Gender and Science* (New Haven: Yale University Press, 1985); Helen Longino, *Science as Social Knowledge* (Princeton: Princeton University Press, 1990); Sandra Harding, *Whose Science? Whose Knowledge?: Thinking from Women's Lives* (Ithaca, NY: Cornell University Press, 1991); Evelyn Fox Keller and Helen E. Longino, eds., *Feminism and Science* (Oxford: Oxford University Press, 1996); Sandra Harding, *Science and Social Inequality: Feminist and Postcolonial Issues* (Urbana and Chicago: University of Illinois Press, 2006); Alison Wylie, "The Feminism Question in Science: What Does It Mean to 'Do Social Science as a Feminist'?", in *The Handbook of Feminist Research*, ed. Sharlene Hesse-Biber (New York: Sage, 2007); Harold Kincaid, John Dupré, and Alison Wylie, eds., *Value-Free Science? Ideals and Illusions* (Oxford University Press, 2007); Heather Douglas, *Science, Policy, and the Value Free Ideal* (Pittsburgh, PA: Pittsburgh University Press, 2009).

² Alison M. Jaggar, "Love and Knowledge: Emotion in Feminist Epistemology," *Inquiry* 32, n. 2 (1989): 151-176. Jaggar writes on the role of social values in science: "These values are implicit in the identification of the problems that are considered worthy of investigation, in the selection of hypothesis that are considered worthy of testing and in the solutions to the problems that are considered worthy of acceptance" (162).

³ Hans Vaihinger, *The Philosophy of "As If": A System of the Theoretical, Practical, and Religious Fictions of Mankind*, trans. C.K. Ogden (New York: Harcourt, Brace, and Co., 1924).

⁴ For further reading on the value of multiple and often contradictory models, and the role of idealizations, see, e.g.: David Lewis, "Logic for Equivocators," *Noûs* 16, no. 3 (1982): 431-441; Nancy Cartwright, *How the Laws of Physics Lie* (Oxford: Oxford University Press, 1983); Richard Levins, "The Strategy of Model Building in Population Biology," in *Conceptual Issues in Evolutionary Biology*, ed. Elliott Sober (Cambridge, MA: MIT Press, 1984); Nancy Cartwright, *The Dappled World* (Cambridge: Cambridge University Press, 1999).

research, of course, does not merely reflect one's values but also the kind of conceptual schema(s) in which one is embedded, which in turn impacts the development of hypothesis and what one accepts as evidence. That is, a particular gender schema impacts how one researches the phenomenon of gender, the kind of model one develops, the purpose of the model, and the scope in which the model is applicable. Moreover, one's conceptual schema(s) impacts how one perceives the phenomenon under investigation in better or worse ways, which not only impacts how one develops a model but the extent to which the model is deemed useful in fulfilling its purpose.⁵

In this way, models for shared intention, and, by extension, joint action, will always reflect some set of values and conceptual schema(s) because those models are intended to be useful within that context. It would not make sense to model what shared intentions might look like for an alien civilization if the purpose is to understand and manage the way humans share intentions. In this case, our models for shared intention already reflect the way the world is for us. However, the further issue is that 'us' is not unified. We do not all share some fundamental conceptual schemas, and thus our sense of our social world(s) including ourselves within them affects how we model features of our social world and for what purpose. This means that models for shared intention discussed in Chapters II and III already reflect more particular values and purposes, even if the purpose of the model is intended to capture universal features of shared intention and joint action. Thus, what feminist philosophers of science and feminist epistemologies have emphasized is that there is an important relationship between the researcher and the research, the knower and the known, and the creator of the model and the model itself.

⁵ Robert C. Richardson, "Biology and Ideology: The Interpenetration of Science and Values," *Philosophy of Science* 51, 3 (1984): 396-420. Richardson argues that certain hypothesis will be uncritically accepted given the prevailing ideology. For example, Richardson argues that the hypothesis that intelligence is genetically determined was uncritically accepted due to the prevalence of biologically based racism.

The reasons to interrogate this relationship are two-fold: (1) it provides us with a better understanding of how bias functions in research and how to minimize its impact in cases where doing so is needed; and (2) it provides us with a better understanding of the purpose of the model and thus the scope of the model's applicability.

My claim thus far is that the set of idealizations currently used within the models offered by Michael Bratman,⁶ Margaret Gilbert,⁷ and Christopher Kutz⁸ cannot account for the conceptual schemas in which we are embedded and their relationship to structural forms of oppression, and thus provide us with models that cannot accommodate forms of joint action that require participants to directly address differences in conceptual schemas in order to jointly resist oppression. Because these models do not provide an understanding of such joint actions, the function of these models as a form of second-order social reality management risks imposing values and conceptual schemas that reinforce forms of oppression. That is, if these models dictate how we ought to share intentions or engage in joint actions, then we risk foreclosing on possibilities for shared intention and joint action – that is, by dictating what they are, we fail to see what they also could be.

In the following, I analyze how the predominate models for shared intention employ idealizations that: (1) increase the likelihood that ethically concerning forms of ideological bias are present in the model; and (2) why that is concerning for understanding both the possibilities of shared intentions and joint action as well as managing this social phenomenon. The layout of this chapter is as follows: Section II offers a brief overview of feminist epistemology and

⁶ Michael E. Bratman, *Shared Agency: A Planning Theory of Acting Together* (Oxford: Oxford University Press, 2013).

⁷ Margaret Gilbert, *Joint Commitment: How We Make the Social World* (Oxford: Oxford University Press, 2014).

⁸ Christopher Kutz, *Complicity: Ethics and Law for a Collective Age* (Cambridge, UK: Cambridge University Press, 2000).

intersectionality as a methodology, epistemology, and theory of oppression in preparation for the case studies. Section III takes up two real world cases of potential joint actions, examining the terms of their success or failure. My point here is to show that due to their idealizations, the three models cannot adequately address these cases or the models would label certain forms of failure as success.

Section II. Epistemology and Intersectionality

Feminist epistemology arises in conjunction with concerns raised by feminist philosophers of science regarding the relationship between the subject engaged in research and the object of research. This epistemological tradition argues that in a “S knows *that p*” model for knowledge, there is connection between *S*, as a knower, and *p*, as knowledge.⁹ The claim that the subject, *S*, has bearing not only on the object, *p*, but also on the verb, *knowing*. Traditional epistemology generally does not concern itself with who the knower is and what bearing that has on object of knowledge as well as how that person arrived at that knowledge. On this model, to ‘know *that p*’ is to claim that *p* can be discovered by any rational agent, given the appropriate training. As Donna Haraway argues, following in the Enlightenment tradition, *S* is supposed to represent an idealized agent who can perform “the god-trick of seeing everything from nowhere.”¹⁰ The claim is thus that traditional epistemology idealizes knowers by excluding concerns of bias and other forms of cognitive error as well as the potential affect that one’s social location has on both *what* and *how* one knows. In this way, traditional epistemology relies on the

⁹ Lorraine Code, “Is the Sex of the Knower Epistemologically Significant?” *Metaphilosophy* 12, no. 3–4 (July 1981): 267–76.

¹⁰ Donna Haraway, “Situated Knowledges: The Science Question in Feminism and the Privilege of Partial Perspective.” in *The Feminist Standpoint Theory Reader: Intellectual and Political Controversies*, ed. Sandra Harding (New York: Routledge, 2004), 86.

idealization of agent-neutrality under the assumption that agent-neutrality is required for impartiality and that impartiality ensures knowledge is neutral (i.e., not tainted by ideology). Feminist epistemology thus focuses on why a methodology that values neutrality is not actually neutral in its values. As Sandra Harding argues, “The more value-neutral a conceptual framework appears, the more likely it is to advance the homogenous interests of dominate groups, and the less likely it is to be able to detect importance actualities of social relations.”¹¹

Both Harding and Haraway argue that because knowers are embedded in the world, a knower or group of knowers cannot offer a complete picture of the world but potentially can offer partial, localized knowledge claims that contribute to a more complete picture of the phenomenon in question. To focus on how knowers (and, by extension, knowledge) are situated or embodied is not to argue that knowledge is wholly relative but rather a claim about how a greater diversity in partial knowledge claims about a phenomenon can result in a better model or models for it.¹² A methodology that does not begin from the idealization of agent-neutrality not only provides better models, but is also power-sensitive, ensuring that we are held accountable for our practices of knowledge production.¹³ This is similar to Kwame Anthony Appiah’s claim that cognitive limitations of human make a grand theory of everything impossible.¹⁴ Appiah argues that, at best, we will have many models for a phenomenon, which need not be compatible. Each model gives us a partial understanding, and combining models can provide a more comprehensive understanding.

¹¹ Sandra Harding, “Introduction: Standpoint Theory as a Site of Political, Philosophic, and Scientific Debate,” in *The Feminist Standpoint Theory Reader: Intellectual and Political Controversies*, ed. Sandra Harding (New York: Routledge, 2004), 6.

¹² Harding, *Whose Science? Whose Knowledge?*.

¹³ Donna Haraway, “Situated Knowledges,” 81-101.

¹⁴ Kwame Anthony Appiah, *As If: Idealization and Ideals* (Harvard: Harvard University Press, 2017).

The tradition of feminist standpoint theory in particular connects this rejection of presumed neutrality to the role that structural oppression plays both in the production of knowledge and for the producers of knowledge.¹⁵ In its broadest formulation, standpoint theory consists in the following claims: (1) knowledge is produced from knowers who are embedded in the world and (2) the production of knowledge is necessarily entangled with one's social-structural positioning. However, there is no automatic process by which peoples' experiences in being members of structurally oppressed groups turns into knowledge about the functions and mechanisms of oppression. Nancy Hartstock argues that the development of a standpoint is an achievement by those who are oppressed, because, given the range and impact of dominate narratives, producing a different narrative (as well as showing how the narratives conflict) is a struggle in itself (and thus a form of resistance in itself).¹⁶ Because knowledge is a group-based achievement (insofar as it is not done by a single individual), it is through forms of shared agency and joint action that such knowledge is produced – not prior to it. In a response to Harding, Gail Polhaus argues for a more deeply social interpretation of standpoint theory by noting the dual role of struggle in knowledge production, emphasizing the role of community-building:

...If we consider community, not as an aggregate of persons holding something in common, but rather as constituted by relations among members, struggling together with the community becomes as important as struggling against a particular system or group of people. To struggle-with would involve building

¹⁵ The idea of a 'women's standpoint' in as a critique of scientific practices comes from: Dorothy Smith, "Women's Perspective as a Radical Critique of Sociology," *Sociological Inquiry* 44, no. 1 (1974): 7-13.

¹⁶ Nancy Hartsock, "The Feminist Standpoint: Developing the Ground for a Specifically Feminist Historical Materialism," in *Discovering Reality: Feminist Perspectives on Epistemology, Metaphysics, Methodology, and Philosophy of Science*, ed. Merrill B. Hintikka and Sandra Harding (Dordrecht, Netherlands: Kluwer Academic Publishers, 1983), 283-310.

relations with others by which we may come to know the world and understand one another, that is the project of building knowing communities.”¹⁷

Polhaus emphasizes that the production of knowledge is not simply about the opposition to oppression but also claim that knowledge productive itself requires relations with others – it is necessarily an intersubjective process. Moreover, because knowledge is an achievement through practices of community-building, this also means that these achievements are not global; they must be further disseminated. For example, through forms of feminist consciousness-raising, those who engaged directly in the process achieve a knowledge claim about misogyny. This knowledge claim is thus not immediately available to all those who are subject to misogyny. That said, those who have direct experience of misogyny might be more likely to understand this knowledge claim because it provides a way to conceptualize their experiences. That is, knowledge in this case alters one’s conceptual schema by shifting preexisting concepts (or their relations) or introducing new ones. Those who are not directly subject to misogyny thus might struggle more in understanding. In the case of shared intentions, this means that agents will already have different conceptual schemas and asymmetrical scopes of knowledge. In the case of knowledge about oppression, these asymmetries do not necessarily track an individual’s social-structural positioning but likely do track it to some extent (even if this does not hold for any particular member of a social group).

Standpoint epistemology overlaps with epistemology practices within intersectionality, particularly in the case of Black feminist and decolonial resistance knowledge projects. Resistant knowledge projects are those that: “grapple with the existential question of how individuals and groups who are subordinated within varying systems of power might survive and resist their

¹⁷ Gail Polhaus, “Knowing Communities: An Investigation of Harding’s Standpoint Epistemology,” *Social Epistemology* 16, no. 3 (2002): 291-292.

oppression.”¹⁸ Such resistant knowledge projects contain similar value commitments about the prioritization of direct experience but more adequately carry out the epistemological demand to further contextualize the process of knowledge production. In this way, such resistant knowledge traditions in the case of feminist thought have been critical of a monolithic approach to standpoint theory, particularly the tendency for white, western women to universalize our particular experiences and problems in broad reference to a “women’s standpoint,” which is why Patricia Hill Collins argues that there is a pressing need to understand that standpoints are many, not one.¹⁹

Intersectionality, more broadly, as a theory of oppression, a methodology, and set of epistemological practices offers in this case a valuable approach for thinking through forms of resistant shared agency in particular. In *Intersectionality as Critical Social Theory*, Collins argues for four main premises that an intersectional approach:

1. Race, class, gender, and similar systems of power are interdependent and mutually construct one another.
2. Intersecting power relations produce complex, interdependent social inequalities of race, class, gender, sexuality, nationality, ethnicity, ability, and age.
3. The social location of individuals and groups within intersecting power relation shapes their experiences within and perspectives on the social world.
4. Solving social problems within a given local, regional, national, or global context requires intersectional analyses.²⁰

Claims (1) – (3) function both as a set of epistemological and ontological claims about the nature of oppression, how practices of knowledge production arise within complex systems of power,

¹⁸ Patricia Hill Collins, *Intersectionality as Critical Social Theory* (Durham, NC: Duke University Press, 2019), 88.

¹⁹ Patricia Hill Collins, *Black Feminist Thought: Knowledge, Consciousness, and the Politics of Empowerment* (New York: Routledge, 1991), 21-37.

²⁰ Collins, *Intersectionality as Critical Social Theory*, 44.

how those practices vary between different interpretative communities,²¹ and the ways in which dominate ideologies (i.e., ones that reflect the interests of those who largely benefit from the oppression of others) can not only inhibit the development and force of resistant knowledge projects, but actively suppress such projects. Suppression in this case can range from the literal destruction of stores of knowledge and disruption of intergenerational knowledge transmission to interpersonal forms of testimonial silencing.²²

The ontological theory of oppression here is broadly that forms of oppression do not share the same source but are necessarily intertwined. Systems of power do not operate in a monolithic way at a global level, but there are dynamic intersections that do shape how forms of power play out within local and global contexts. Moreover, social categories – such as race, class, or gender – do not function the same in all contexts or communities, such that analysis of these categories, their relational configurations and their meanings, must develop through a particular context and its history (through which the configurations and meanings developed).²³ For this reason, theorizing oppression and related social categories through what Kimberlé

²¹ Collins, *Intersectionality as Critical Social Theory*, 116-144. Collins' analysis of interpretative communities is valuable in part because she highlights how different interpretative communities have their own norms, and forms of legitimation and justification. This places outsiders to a community at an epistemic disadvantage because their own practices might not be seen as correct and thus their claims treated as suspect or not as a form of knowledge. However, this means, for example, that within a more dominate interpretative framework (e.g., all rational beings can arrive at the same knowledge through the use of reason alone), the oppressed interpretive communities will be treated as not epistemologically rigorous or as not a site of genuine knowledge production. Thus, knowledge claims from members of oppressed groups about the nature of their oppression (i.e., the knowledge claims from their resistant knowledge projects) will be ignored or actively suppressed. This ignorance or suppression often directly appeals to the standards for legitimacy and justification within a dominate interpretive community under the claim that their parochial practices are universal (particularly because they follow from a particular picture of rationality, experience, and access to knowledge).

²² Kristie Dotson, "Tracking Epistemic Violence, Tracking Practices of Silencing," *Hypatia* 26, no. 2 (Spring 2011): 236-257.

²³ Collins, *Intersectionality as Critical Social Theory*, 54-84, 157-188.

Crenshaw calls a “single categorical axis”²⁴ does not provide an adequate ontological understanding of oppression. That is, intersectionality offers a theory of oppression that is more explanatorily useful than theories that treat different forms of oppression as analytically separate or presume that social categories such as ‘woman’ have a universally fixed meaning and/or correlate to a set of universally shared experiences. As Crenshaw writes in the case of monolithic feminist assumptions, the category of ‘woman’ functions similarly to the claim that an agent-neutral approach to knowledge allows particular situated knowers to speak from nowhere and thus present their claims as impartial:

The value of feminist theory to Black women is diminished because it evolves from a white racial context that is seldom acknowledged. Not only are women of color in fact overlooked, but their exclusion is reinforced when *white* women speak for and as *women*. The authoritative universal voice - usually white male subjectivity masquerading as non-racial, non-gendered objectivity - is merely transferred to those who, but for gender, share many of the same cultural, economic, and social characteristics. When feminist theory attempts to describe women's experiences through analyzing patriarchy, sexuality, or separate spheres ideology, it often overlooks the role of race. Feminists thus ignore how their own race functions to mitigate some aspects of sexism and, moreover, how it often privileges them over and contributes to the domination of other women.²⁵

Thus, resistant knowledge projects require attention not only to who is participating in the process of knowledge production and larger communities in which those participants exist (in which, for example, forms of structural ignorance might produce highly distorted or inaccurate claims about oppression), but also to the potential scope to which those knowledge claims apply. Here, the methodological claim is that standpoints can be partial, dynamic models that produce

²⁴ Kimberlé Crenshaw, “Demarginalizing the Intersection of Race and Sex: A Black Feminist Critique of Antidiscrimination Doctrine, Feminist Theory and Antiracist Politics,” *University of Chicago Legal Forum* 140, no. 1 (1989): 140.

²⁵ Crenshaw, “Demarginalizing the Intersection of Race and Sex,” 154.

more explanatorily powerful theories about the particular formations and contours of power in a given case.

These models idealize social categories that can be better or worse, methodologically, epistemologically, *and* ethically. What an intersectional approach highlights more broadly, that those engaging in the process of theorizing can uncritically adopt current forms of ontological oppression and engage in dominate norms or practices, even as they themselves are engaged in projects trying to resist them. In this way, an ontological theory of oppression not only requires a *multi-axis* approach, but shows that the larger ontological theory offered by intersectional theorists could only have come about through a *multi-axis* epistemological approach in the first place (and, in fact, did). Moreover, as Deborah King argues, this multi-axis approach cannot be understood on an *additive* analysis (e.g., race + gender + class) wherein all forms of oppression are seen as essentially the same (and thus broadly interchangeable in the process of analysis), and one can understand intersections by merely adding the effects together.²⁶ Rather, as Collins argues, intersectionality is a lived, communal experience, and this lived experience can shift to greater or lesser extents as people travel between different communities, and this includes travel between different interpretative communities.²⁷

Standpoint epistemology offers the claim that knowledge production about oppression ultimately tracks an embodied experience of that oppression but an intersectional approach offers an enhanced picture this process. By starting from a multi-axis approach, practices of knowledge production can more adequately address the complex, contextual, and relational nature of forms

²⁶ Deborah King, "Multiple Jeopardy, Multiple Consciousness: The Context of a Black Feminist Ideology," *Signs* 14, n. 1 (1988): 42-72.

²⁷ See both: Collins, *Intersectionality as Critical Social Theory*, 126-144; María Lugones, "Playfulness, "World"-Traveling, and Loving Perception," in *Pilgrimages/Peregrinajes: Theorizing Coalition Against Multiple Oppressions* (Lanham, MA: Rowman & Littlefield Publishers, 2003), 77-100.

of social inequality and power, but also emphasizes that the multi-axis approach is not merely about the intersections but rather about the ways in which such intersections exist within forms of unjustified social and political hierarchies.²⁸ In this way, intersectionality as a methodology and epistemological is not a claim about the value of “diversity” for knowledge (as it is sometimes treated), but is, as Collins argues, an *analytical strategy*²⁹ for theorizing about the relationship between hierarchy and intersection. Such projects ought to understand identities in a necessary relation to one another, frame the structuring of identities within historical and current conditions of material inequality, and start from the position that, as Collins writes: “individuals and groups differentially placed within intersecting systems of power have different points of view on their own and others’ experiences with complex social inequalities, typically advancing knowledge projects that reflect their social locations within power relations.”³⁰ These intersections start from the point of race and gender in Crenshaw’s case and more broadly include (but are not limited to) intersections of class, sexuality, disability, immigration/citizenship status, and one’s status within or relationship to historical and ongoing practices of colonialization and colonialism.

The connection between experience, oppression, and knowledge also provides an insight about ignorance. Linda Alcoff traces three varieties of ignorance as they relate to social-structural positionings.³¹ First, ignorance arises in virtue of the situatedness of knowers, which

²⁸ Collins, *Intersectionality as Critical Social Theory*, 49. My wording here is intentional, tracking the claim that intersectionality has six key constructs: power, social inequality, social context, social justice, complexity, and relationality.

²⁹ Patricia Hill Collins, “Intersectionality’s Definitional Dilemmas,” *Annual Review of Sociology* 41, no. 1 (2015), 11.

³⁰ Collins, “Intersectionality’s Definitional Dilemmas,” 14.

³¹ Linda Martín Alcoff, “Epistemologies of Ignorance: Three Types,” in *Race and Epistemologies of Ignorance*, ed. Shannon Sullivan and Nancy Tuana (Albany, NY: SUNY Press, 2007), 39-57.

means that knowers are not epistemically equal in virtue of their situatedness.³² Second, social categories and identities play a particular role in the creation and maintenance of knowledge or ignorance. Group identity is epistemically relevant because: “groups will sometimes operate with different starting belief sets based on their social location and their group-related experiences, and these starting belief sets will inform their epistemic operations such as judging coherence and plausibility.”³³ Third, structural forms of oppression produce ignorance of oppression. Drawing on Charles Mills’ work in “White Ignorance,”³⁴ Alcoff argues this third kind of ignorance results in the broad claim that: individuals in more privileged social-structural positionings are more likely to be predictably ignorant about the forms of structural oppression from which they directly benefit. Charles Mills refers to this ignorance as an “inverted epistemology”: “a particular pattern of localized and global cognitive dysfunction (which are psychologically and socially functional), producing the ironic outcome that whites will in general be unable to understand the world they themselves have made.”³⁵

This ignorance is a function of investment. If you benefit from structural oppression, you are more likely to fail to see the extent of that oppression or see the fact of structural oppression at all. Members of privileged social groups are thus less likely to seek out knowledge about forms of structural oppression from which they benefit. Privilege itself can also place epistemic limitations that prevent the person from understanding that there’s a problem in the first place. If a person does not have a direct, phenomenological experience with a form of oppression, then

³² Lorraine Code, “Taking Subjectivity into Account,” In *Feminist Epistemologies*, ed. Linda Alcoff and Elizabeth Potter (New York: Routledge, 1993), 15-48.

³³ Alcoff, “Epistemologies of Ignorance: Three Types,” 45.

³⁴ Charles W. Mills, “White Ignorance,” in *Race and Epistemologies of Ignorance*, ed. Shannon Sullivan and Nancy Tuana (Albany, NY: SUNY Press, 2007), 13-37.

³⁵ Charles W. Mills, *The Racial Contract* (Ithaca, NY: Cornell University Press, 1997), 18.

they are less likely to know it exists or understand its mechanisms, especially if they are surrounded by others or consume media that reflects this experiential lack as well.

In this way, knowledge claims about oppression produced from multi-axis, oppressed social-structural positionings tend to be more useful for understanding and managing (i.e., responding to) oppression because they are less likely to be subject to the relevant mechanisms of ignorance. Oppressed individuals are less likely to be invested in maintaining the current power relations because, in part, they are less likely to be self-deceived about power relations that are not to their benefit.³⁶ One reason for this type of epistemic advantage is Collins' concept of *outsiders-within*, which states that oppressed individuals often must learn to navigate the dominant worldviews (e.g., bias, assumptions, theoretical frameworks, larger narratives) in order to survive.³⁷ As a result, *outsiders-within* are more likely to notice how the actual experiences of non-dominant groups conflict with dominant representations, and are thus in a better position to both critique such dominant representations and generate more accurate knowledge about their lives.

However, a standpoint approach does not romanticize or idealize members of oppressed groups nor does it ignore the very real impacts of oppression on members of oppressed groups. Similar to Collins' concept of outsiders-within, Uma Narayan argues that the epistemic advantage that members of oppressed groups have is a kind of "double vision" borne from the necessity of having to navigate multiple contexts and practices – both dominate and non-

³⁶ Harding, *Whose Science? Whose Knowledge?*, 138-42.

³⁷ Collins, *Black Feminist Thought*, 11-13, 16, 36, 94, 204, 232-33; Patricia Hill Collins, "Learning from the Outsider Within: The Sociological Significance of Black Feminist Thought," *Social Problems* 33, no. 6 (1986) :14-32.

dominate.³⁸ She argues that there is a dark side to this double vision. First, double vision does not necessarily produce (what Chela Sandoval refers to as) *differential oppositional consciousness*, which plays an important role in the development of resistant knowledge projects.³⁹ Rather, Narayan argues that this double vision might beget assimilation to the dominate worldview (i.e., dominate conceptual schemas and/or ideology) or other forms of internalized oppression. Thus, it does not follow from standpoint theory that we can idealize the capacities of oppressed people in modeling the mechanisms of knowledge production in this way, especially when damage caused by oppression (particularly psychic damage) and realities of survival work together to create a loss of intergenerational, non-dominate knowledge, or prevent individuals from adequately conceptualizing the role of forms of oppression in their lives. Second, double vision does not necessarily lead to a unified experience or understanding, which prevents certain kinds of knowledge production because people compartmentalize between different contexts. That is, having to navigate multiple and contradictory conceptual schemas within one's everyday life can reduce one's "fluency" in contexts in which they navigate.⁴⁰ Thus, if standpoints represent different models for forms of structural oppression, then the claim that standpoints rely on a form of embedded epistemic advantage should not romanticize the actual conditions that produce this double vision nor should it result in putting all members of oppressed groups on an epistemic pedestal.

³⁸ Uma Narayan, "The Project of Feminist Epistemology: Perspectives from a Nonwestern Feminist," in *The Feminist Standpoint Theory Reader: Intellectual and Political Controversies*, ed. Sandra Harding (New York: Routledge, 2004), 221.

³⁹ Chela Sandoval, "U.S. Third World Feminism: The Theory and Method of Oppositional Consciousness in the Postmodern World," *Genders* 10 (1991): 1-24.

⁴⁰ Narayan compares this to cases of multilingualism in which the person is less fluent in their languages compared to monolingual speakers of each language.

Of course, one common concern raised about standpoint theory is whether and to what extent people can overcome such epistemic limitations. If such limitations cannot be overcome at all, then this radically shuts down the claim that shared intentions are common. However, it seems that people can overcome such limitations, but the question is to what extent and in what ways. Thus, I agree with Narayan's point that just because knowledge is socially constructed, it does not follow that those in different and/or more privileged social-structural positionings cannot gain an understanding or sympathy with forms of oppression they do not endure.⁴¹ However, Narayan argues that this potential for understanding and sympathy also highlights the possibility of failure:

Those who display sympathy as outsiders often fail both to understand fully the emotional complexities of living as a member of an oppressed group and to carry what they have learned and understood about one situation in the way they perceive another. It is commonplace that even sympathetic men will often fail to perceive subtle instances of sexist behavior or discourse.⁴²

In this way, the process of understanding is also marked by failure, or, in other words, displays of ignorance. This is also why members of oppressed groups have good reason to question forms of sympathy or attempts at solidarity from those outside the group who do not endure the form(s) of oppression in question and/or actively benefit from the existence of that oppression. As Narayan argues:

On the one hand, one cannot but be angry as those who minimize, ignore, or dismiss the pain and conflict that racism and sexism inflict on their victims. On the other hand, living in a state of siege also necessarily makes us suspicious of expressions of concern and support from those who do not live these oppressions. We are suspicious of the motives of our sympathizers or the extent of their sincerity, and we worry, often with good reasons, that they may claim that their

⁴¹ Narayan, "The Project of Feminist Epistemology."

⁴² Narayan, "The Project of Feminist Epistemology," 220.

interest provides a warrant for them to speak for us, as dominate groups throughout history have spoken for the dominated.⁴³

For the same reason that choosing an object of study is not conceptually or evaluatively neutral, one's interest in the oppression of others is likewise not neutral. Even in cases where one's interest reflects values about the wrongfulness of structural oppression, one's interest does not necessarily guarantee that one understands the relevant and different conceptual schemas in play, resulting in behavior that actually reproduces the mechanisms of oppression in the attempt to resist them.⁴⁴ For instance, a person's interest might reflect their desire to position themselves as a savior for oppressed groups of which they are not members. In such cases, a person might genuinely believe they are contributing to or, worse, creating the resistance struggles of another group. This problem is often reflected in the idea that one can simply drop-in to a different community, devoid of any knowledge about that community and the people who comprise it, and address ongoing problems better than long-term members of a community can.⁴⁵ Such cases do not merely represent a failure to share intentions but can actively harm ongoing forms of shared agency that members of the communities have developed amongst themselves.

⁴³ Narayan, "The Project of Feminist Epistemology," 219.

⁴⁴ In this way, interest in oppression from which one benefits also takes up interests of the dominate group by functioning as a form of oppressive curiosity. That is, one might assume their interest is counter to the interests of the dominate group in which they exist but, in all actuality, is complementary to it. One might also point out that the assumption that one has the right to know or access the details of resistant knowledge projects is itself a colonial attitude, and, even if unintended, access to such knowledge by members of dominate groups risks being weaponized in line with group-level interests in maintaining domination.

⁴⁵ This 'drop-in' mentality often correlates to a 'drop-out' mentality; that is, because the person is not subject to the form of oppression, they are likely to feel more at ease by dropping-out of ongoing organizing because it has little to no bearing on the course of their life. But this 'drop-in, drop-out' mentality is already different from those who are already members of the community and cannot 'drop-out' of being oppressed (and thus will have a harder time abandoning both ongoing organizing efforts and the communities in which they are embedded). Given that anti-oppression resistance, or movement building, and the knowledge needed for resistance is situated, 'drop-in, drop-out' participants create problems for the success of both shared intentions and joint actions.

Broadly, then, the value of standpoint theory combined with the methodological and epistemological framework of intersectionality is to emphasize the relationship between: (1) a person's or group's investment in harmful mechanisms; (2) a person's or group's direct experience of harm; (3) the role of investment and experience in both the production of knowledge about oppression and the production of ignorance about oppression; (4) the need to prioritize resistant knowledge projects borne from multi-axis experiences of subordination and/or domination, in part for reasons related to accuracy; and (5) to prioritize such knowledge projects without romanticizing (or, worse, fetishizing) members of oppressed groups as knowers or the conditions under which such knowledge arises (or the fact that such resistance knowledge projects are necessary in the first place). The normative intuition that follows is that we ought to give epistemic priority to such standpoints in cases of disagreement about structural oppression, especially to multi-axis standpoints.

In preparation for the case studies in Section III, I would like to draw out a few claims about the relationship between this literature and literature on shared intentions. First, people are situated within different and multiple conceptual schemas, which affects how those who are differently situated successfully share intentions and carry out joint actions. Second, oppressed groups are more likely to navigate multiple conceptual schemas that clash in significant ways *and* this clash happens in occurrence with related mechanisms of oppression, such that this clash can be understood between dominate and non-dominate conceptual schemas and between interpretive communities.⁴⁶ Third, resistant knowledge projects can shift both dominate and non-dominate conceptual schemas. This shift can be the introduction of new concepts or conceptual

⁴⁶ This is not to say that there is one dominate conceptual schema but rather that the dominate/non-dominate divide does not exist independent of more particular contexts.

frameworks, criticism of prevailing conceptual schemas (particularly ones that create and/or reproduce forms of oppression), a reorientation of a schema, or the introduction of a paradigm-changing metaphor.⁴⁷ Fourth, the relationship between ignorance, knowledge, and oppression requires attention to how one's perception that an intention is shared can be wrong. For example, when differently situated agents participate anti-oppression movements, failures often result from a push to oversimplify highly complex mechanisms of oppression by attempting to isolate one form from another. These failures occur both from those who attempt to do work in anti-oppression movements of which they are not the subject (e.g., heterosexual and cis-gender people within queer liberation movements) *and* from those who hold more privileged social-structural positioning within a movement of which they are the subject (e.g., white women within feminist movements).

For shared intentions in particular, this means that those subject to multiple forms of oppression might not be in the position to share intentions with those engaged in this kind of oversimplification because an unequal compromise serves as the basis for action, not genuinely shared activity. This compromise is unequal because certain groups, in virtue of their social-structural positioning, are more likely to be placed in the position of having to compromise, especially for reasons related to survival. Due to the mechanisms of ignorance, those with more privileged social-structural positionings are less likely to know that this kind of compromise is happening and more likely to mistake it for genuinely shared activity, or more likely to believe that this compromise is justified in order to be effective for a single-axis movement (and thus more likely to demand this from those who exist at multiple axes). In this way, all participating agents are not genuinely sharing intentions but rather different sub-sets of participants are

⁴⁷ Collins, *Intersectionality as Critical Social Theory*, 41-45.

sharing different sets of shared intentions. The relationship between the sets of shared intentions is strategic coordination but this is only clear to those who are having to compromise. Because the role of compromise is not epistemically available to all agents, further failures result from mistaken beliefs as to what extent the activity is shared, which impacts future development of goals and plans for action within a movement.

Section III. Case Studies

The following two cases each capture different but related aspects of why the idealizations used by Bratman, Gilbert, and Kutz are of methodological, ethical, and ideological concern. Case 1 focuses on a potential joint action in which participants diverge as to whether it was successful, but do agree that a certain form of failure occurred. Case 2 focuses more on how conceptual differences in the case of terminology impact the development of a successfully shared intention and joint action.

Case 1: Building a Community Together

In “The Costs of Exclusionary Practices in Women’s Studies,” Maxine Zinn et al. argue that women of color and working-class women have been largely excluded by feminist theory because of its institutionalization within academia.⁴⁸ Accessing academic institutions is easier for those with both racial and class privilege, which thus largely excludes members of such groups from vital roles in developing feminist theory.⁴⁹ As a result, academic feminist theory often

⁴⁸ Maxine Baca Zinn, Lynn Weber Cannon, Elizabeth Higginbotham, and Bonnie Thornton Dill, “The Costs of Exclusionary Practices in Women’s Studies,” *Signs* 11, n. 21 (1986): 290-303. This article was partially a response to the exclusionary practices of *Signs* in particular.

⁴⁹ This is not to say that feminist theory only comes from academics. Rather, it is to point out that ‘academic feminist theory’ is often legitimated as *the* source of feminist theory, and thus excluding and/or limiting the participation of certain groups of women.

reflects the experiences and interests white and middle- to upper-class women specifically because such women were able to set the terms of the discussion within the literature, given that they are more likely to be able to access and secure academic positions in the first place. As an example, the authors refer to a *Signs*-sponsored conference at Stanford University in 1983 on the subject of ‘communities of women’ in which the organizers predetermined how the theme ought to be interpreted by potential presenters.

I use this example for three reasons. First, it captures well the problems with the three standard idealizations: an idealized epistemology, an idealized social ontology, and the idealization of agent-neutrality. Second, given that communities are the sites through which knowledge production occurs, the example offers a multi-tiered analysis of how non-shared frameworks for community and community-building (that themselves arise in dialectical response to systems of power) shape how people try to build new forms of community specifically geared to the production of resistant forms of knowledge. In this case, one parochial notion of community and the invitation to join displays how the conditions for and practices of one community can both: (1) flourish as a direct result of oppression or exclusion of other communities and frameworks for community; and (2) display an ignorance about the conditions of their community by assuming it offers the best starting point for new forms of community-building, particularly with those excluded, or, worse, a ‘neutral’ starting point through which to reflect on and build community together. Lastly, and third, because it offers a multi-tiered, or meta-reflection, on the relationship between community-building as knowledge production and differences in communal practices in how knowledge production occurs (as well as who and what that knowledge is for), it highlights why theorizing shared agency, which is a vital part of these processes, necessarily requires a power-sensitive and contextual analysis that actively

addresses the complex ways in which forms of oppression affect these processes, for better or for worse, and thus by extension the ability for unequally situated agents to engage in these processes together.

This example thus starts with the attempt to develop a shared conference theme in which the organizers began by using a parochial notion of community and send out invitations on those terms:

In 1981, the planners of a conference on communities of women asked Elizabeth Higginbotham to submit an abstract for a paper. The expectation communicated in the letter of the invitation was that her research would demonstrate the applicability to Black women of a concept of women's communities set forth by white feminists. Instead of attempting to alter her work to fit such a model, Higginbotham wrote to the organizers and challenged their narrow definition of communities of women. Higginbotham noted that, unlike their white sisters who are often excluded from male-dominated spheres or retreat from them, the majority of Black women are ordinarily full participants in mix-sex spheres and make unique contributions to both the definitions and problems.⁵⁰

In the editorial foreword to the conference proceedings in *Signs*, the organizers reflect on what they take to be the failures and successes of the conference.⁵¹ First, they note that the conference invitation asked potential participants to reflect on the topic – communities of women – in a way that presupposed the values and experiences of particular groups of women:

We did not realize, until courteously yet explicitly advised of our failure of insight, that this focus might be seen to limit the topic's appropriateness to the interests of middle-class white women in their struggle against the dominance of white men.⁵²

⁵⁰ Zinn et al., "The Costs of Exclusionary Practices in Women's Studies," 295.

⁵¹ "Editorial," *Signs* 10, no. 4 (1985): 633-636. The authors' names are not listed but *Signs* had a communal editorial staff at the time. However, the editor-in-chief at *Signs* from 1980-1985 was Barbara Gelpi. Based on an interview with Gelpi in 2013, the organizers (and thus the authors of this editorial) likely include, but are not necessarily limited to: Barbara Gelpi, Myra Strober, Carol Jacklin, Susan Johnson, Shelly Rosaldo, Margery Wolf, and Estrelle Freedman. See: Barbara Charlesworth Gelpi, interview by Allison Tracy, March 12, 2013, interview part 2 (57-73), transcript, Stanford Historical Society Oral History Program (SCO932), Department of Special Collections and University Archives, Stanford University Libraries and Academic Information Resources, Stanford, CA, <https://purl.stanford.edu/mn590gg7877>.

⁵² "Editorial," 634.

Thus, they understand their failure to be largely epistemic.⁵³ They describe this failure in terms of their own epistemic limitations; that is, their own conceptual framework for modeling ‘community’ and its potential for achieving feminist goals. More pressing, this epistemic failure highlights the way in which their model was borne from and reproduced a form of ignorance about formations of community that were not captured by their model.

However, the editors argue that these failures actually made it possible for the conference to succeed in its goal:

Yet if conference planning and conference papers alike demonstrated the problematic nature of women's communities, the conference itself was, however paradoxically, the celebration of community that had been our first wish... The conference succeeded in being what it was discussing – a community – only through the help and participation of a great many people.⁵⁴

Nevertheless, it is not clear that all participants understood the conference to be a success, even in this way. This is evidenced by the example later being cited as both a foundational and an epistemic failure. If the failure in this case were merely a local and easily remediable epistemic failure on the part of the organizers, then it might be easier to motivate the claim that the conference was successful in achieving its goal. However, Zinn et al. argue that the problem is also structural, or institutional (i.e., foundational): those with easier access to academia get to set the terms of the discussion, which shapes how those very same people understood the terms of

⁵³ Barbara Charlesworth Gelpi interview, 73. Gelpi notes that the communal editorial staff at *Signs* was “very inclusive” for those involved but also added when discussing the relationship between the all-white editorial staff/organizing committee and the conference: “But you could see that that group had in it no women of color, so that’s significant. Now, we did hear [about issues of race], we could see that, but we weren’t then, ourselves.” This portion of the interview makes it clearer that the notion of community used in the conference theme was, in part, based on their experiences in an all-white editorial group.

⁵⁴ “Editorial,” 635-636. Moreover, in her 2013 interview, Gelpi comments further on the conference as successful.

community and the invitation to be included in the community of another. As bell hooks captures the dynamic of this problem well:

One reason white women active in the feminist movement were unwilling to confront racism was their arrogant assumption that their call for Sisterhood was a non-racist gesture. Many white women have said to me, “we wanted black women and other non-white women to join the movement,” totally unaware of their perception that they somehow “own” the movement, that they are the “hosts” inviting us as “guests.”⁵⁵

The problem is not merely the epistemic failure of the organizers, but also a failure to understand how racial and class privilege had already allowed white and economically privileged women within academia to set the terms of the discussion, such that revision to the conference goal under the guise of ‘inclusion’ continued to center their preset terms, even in focusing on the failure of these preset terms. In this sense, they failed to understand that the very possibility of their community functioned as an exclusion, such that their notion of community could not be merely modified in line with the epistemic criticism. As Collins argues, this white-centric feminism itself already had advantages of obtaining academic legitimacy in the first place:

The sizeable number of women students and faculty who entered the academy and who supported women’s studies programs provided a visible group of advocates for feminist scholarship. Moreover, women students were often white and middle-class, a status that highlighted the significance of gender. These women brought educational and financial resources with them, and they also could recruit men from their social circles as potential allies.⁵⁶

This not only also further pre-set the terms of the discussion but also highlights how notions of community, access, and exclusion resulted in earlier forms of academic feminism to broadly take a single-axis approach in the case of gender. What’s also noteworthy about Collins’ point is that white women’s sense of community might not have explicitly included white men but, at

⁵⁵ bell hooks, “Sisterhood: Political Solidarity Between Women,” in *Feminist Theory: From Margin to Center* (Boston, MA: South End Press, 1984), 53.

⁵⁶ Collins, *Intersectionality as Critical Social Theory*, 101.

minimum, relied on alliances with white men in order for their form of community to be possible. In this sense, the difference is not about the inclusion of men but how men are included and why. In this case, epistemic failure is not local to the conception of community put forth in the invitation but rather the resulting structural inequalities both within and outside of academia, including what it means for some to travel between academic and non-academic worlds, especially in the case of resistant knowledge projects.

The organizers argued that despite their failures, criticism of the conference's goal was the reason they successfully met the goal, which was the creation of a community that reflected on the topic 'communities of women.' Zinn et al.'s use of the example comes later, and in response to ongoing problems in women's studies (and *Signs* in particular), and, as I take it, likely reflects a disagreement both about the terms of failure and success for this conference. In this case, one might argue that conference might have been more akin to a form of compromise, but, in labeling it a success, the organizers engage in further epistemic failures that reinforced the initial problem.

What this case also highlights are the unequal ways people travel between different interpretative communities. Collins' analysis of interpretative communities is valuable here in part because she highlights how different interpretative communities have their own norms and forms of legitimation and justification. This places outsiders to a community at an epistemic disadvantage because their own practices might not be seen as correct and thus their claims treated as suspect or not as a form of knowledge. This is true in the case of academic and non-academic interpretative communities broadly, but, in this case, it specifically places practices of knowledge of production arising largely outside of academic spaces at a disadvantage because of the ways academic forms of knowledge production are often treated as more legitimate. This

means, for example, in a more dominate interpretative framework (e.g., all rational beings can arrive at the same knowledge through the use of reason alone) oppressed interpretive communities will be treated as not epistemologically rigorous or as not a site of knowledge production entirely. Thus, knowledge claims from members of oppressed groups about the nature of their oppression (i.e., the knowledge claims from their resistant knowledge projects) will be ignored or actively suppressed. This ignorance or suppression often directly appeals to the standards for legitimacy and justification within a dominate interpretive community under the claim that their parochial practices are universal (particularly because they follow from a particular picture of rationality, experience, and access to knowledge).

When Zinn et al. note that Higginbotham was asked to interpret the relevance of white feminist notions of community for Black women, this places a priority upon and assumes the legitimacy of those parochial notions of community over and above various Black feminist practices of community in general. Moreover, the organizers failed in their demand for a particular kind of traveling. In “Playfulness, “World”-Traveling, and Loving Perception,”⁵⁷ María Lugones argues that travel between different worlds (or interpretative communities) and the demands of world-traveling do not play out equally. Because these worlds often track our own social-structural positionings (and our traveling between worlds can also shift those positionings contextually), those who inhabit more relatively privileged positionings within a particular domain will often have little reason to travel to non-dominate worlds, while those who inhabit more oppressed positioning will likely be forced to travel between dominate and non-dominate worlds. The demand in this case is placed on those invited to the conference to travel to a more dominate interpretative community allowing those within that dominate world to remain

⁵⁷ María Lugones, “Playfulness, “World”-Traveling, and Loving Perception,” 77-100.

safe (or largely unchallenged) in their positioning. Because this traveling is also about community-specific practices of knowledge production, justification, and legitimation, the demand to travel, in this case, is the demand to apply a dominate world (or interpretative community) to a non-dominate one, assuming uncritically that this application is necessarily to the benefit of those being asked to travel. Moreover, the assumption that different practices can easily translate glosses over the extent to which interpretative communities can significantly differ in their practices, which might affect intelligibility between one world and another. In addition, the notion that people ought to try translate non-dominate practices to dominate ones assumes that certain forms of knowledge ought to be translated in the first place, that all knowledge is open access rather than sometimes community-specific (and that such communities might actively resist the larger dissemination of their practices or knowledge claims).⁵⁸

What this case valuably highlights is the differences in the kinds of joint actions aimed at resisting oppression. Accounts of joint action, if they are to be widely applicable should also be able to capture how shared processes of deliberation occur in figuring how best to engage in forms of epistemic resistance together is already a form of joint action. That is, feminist theory, like other forms of feminist political action, is a joint undertaking, which thus raises concerns

⁵⁸ I note here concerns about de-radicalization and misuses of intersectionality academic contexts as one such case. See: Sirma Bilge, "Intersectionality Undone: Saving Intersectionality from Feminist Intersectionality Studies," *Du Bois Review* 10, no. 2 (2013): 405-425; Sirma Bilge, "Whitening Intersectionality: Evanescence of Race in Intersectionality Scholarship," in *Racism and Sociology: Racism Analysis Yearbook* (Berlin: Lit Verlag/Routledge, 2014), 175-205; Collins, *Intersectionality as Critical Social Theory*, 123-126. Also, the notions of community control over community-specific knowledge, especially in the case of resistant knowledge projects, have become more pressing given contemporary intersections between Big Tech, search engines, and academic institutions. See: Safiya Umoja Noble, *Algorithms of Oppression: How Search Engines Reinforce Racism* (New York: New York University Press, 2018), 129-133. As Noble writes on mass digitization of information: "... the ethical considerations that arise when participants in marginalized communities are unable to participate in the decision making of having content they create circulate to a far wider, and outsider, audience. These are the kinds of issues facing information workers, from the digitization of indigenous knowledge from all corners of the earth that are not intended for mass public consumption, to individual representations that move beyond the control of the subject. We cannot ignore the long-term consequences of what it means to have everything subject to public scrutiny, out of context, out of control" (132).

about who the ‘we’ of the joint action is, and who is excluded or only provisionally included in this ‘we.’ In this case, the ‘we’ started with the organizers, who invited potential participants, which already frames, as hooks notes, ‘we’ in terms of the founders, not those they invited. However, the question is then whether those who accepted the invitation shared the intended goal for the conference by accepting and participating, or if this not insufficient (i.e., given that there are those who accepted the invitation in order to critique the goal of the conference). This is not a trivial question because it has bearing on the terms of success.

For this reason, this case highlights key questions that the predominate models of shared intention cannot sufficiently answer. While the three models all offer a narrow set of success conditions for a shared intention, there is still a question of who determines the failure or success of a joint action, which functions independently of whether the participants successfully shared an intention. However, such predetermined success conditions prevent us from asking a more fundamental question, which is: who determines whether an intention has been successfully shared? For Kutz and Gilbert, those who accepted the invitation seem to meet the conditions for sharing the intention, even though some of those who accepted took issue with the framing of the goal. Relying on Kutz’s account, this is not an issue because the intention to participate does not require one to intend the goal of the action. Gilbert, on the other hand, would likely take the acceptance of the invitation as an expression of personal readiness. Gilbert’s allowance for coercion raises further concerns because disagreement about the goal by invited participants is irrelevant. Following hooks’ analysis, if the terms of academic feminist theory are preset by those with racial and class privilege (as well as the privilege of institutional prestige, i.e., Stanford), then it creates a coercive dynamic on those without racial or class privilege by

presetting the terms of participation. That is, even for those who want to critique this conference or the journal *Signs*, the mechanism for critique often occurs at the very site of exclusion.⁵⁹

In this way, both Kutz and Gilbert might argue that there was a successfully shared intention, even if there is reason to believe that not all involved intended the goal of the action or disagreed with the goal, which is deeply concerning given that the goal was about the role of community in resisting oppression. That is, one would be committed to saying, in broader terms, that certain participants took resisting their oppression as the goal, even though the framing of the goal reinforced the very terms of that oppression. On the other hand, one might argue that since both Gilbert and Kutz require a shared background between participants – either conceptual (Kutz) or perceptual (Gilbert), then there was not shared or participatory intention in this case, but that only serves to highlight how their idealizations make their models largely inapplicable.

This case also raises the question of what the relationship between failure and success is. Based off the models in question, the relationship functions as follows: failure to share an intention (or the absence of participatory intention) means that there is no joint action, but a successfully shared intention does not guarantee a successful joint action (in which ‘success’ is determined by the goal of the action); however, it does seem to follow that a successful shared intention will result in an successful attempt to act jointly (in which ‘success’ is about the attempt itself, not whether the attempt was successful). The conference organizers argue that their failures were a determinate of the success of the conference. While in this case there is reason to believe that the conference likely was not successful in its goal due to both individual and structural failures, it does raise a more general question about failures of individual

⁵⁹ In this case, this applies to the initial critique of the conference, critique of participations at the conference, and the later critique in *Signs* itself.

participants and the success of a joint action, and whether a failure can be a condition for success. Kutz's model, for instance, allows for the failure of individual participants without undermining the success of the joint action, but Bratman and Gilbert do not. However, Kutz's allowance runs him into a different problem, namely that determining the difference between individual participants who fail in light of their participatory intentions and non-participants, who wrongfully see them themselves as participants.⁶⁰ That said, one might note that the ability to successfully share intentions in certain cases only happens in virtue being able to critically reflect on the failure of previous attempts, and necessarily so. For instance, if you and I fail to share an intention due to a clash in our conceptual schemas that both of us failed to notice in virtue of our immersion in different conceptual schemas, then our ability to successfully share an intention can only happen in virtue of that failure because the failure itself illuminates what was previously obscured to both of us: namely, the clash in our schemas and the nature of that clash. Of course, this presents the failure as equal between both agents. In the conference case, the failure was one-sided. In order for the organizers to see their failure, the potential participants had to point it out, which potentially constitutes a form of *epistemic exploitation* (e.g., the demand placed on members of oppressed groups to educate those in privileged groups about their own oppression).⁶¹ In this way, the failure (as well as any ongoing related failures) happens

⁶⁰ I take this to be a general problem for Kutz. If anyone who holds a participatory intention is, in fact, a participant in the collective action, then this creates practical problems for his account of complicity, because the scope of who is complicit seems unreasonably largely. Further, one might argue that a sense of participation is not always the same as actually participating, even if the participation is poor or undermines the action. For instance, if I think, "I intend to participate (or am participating) in this dance contest," it is relevant as to whether I am actually taken by others to be a contestant in the contest and not just some random person who decided to jump on stage halfway through and crash the party.

⁶¹ For further discussion, see: Nora Berenstain, "Epistemic Exploitation," *Ergo* 3, no. 22 (2016): 569-590; Audre Lorde, "The Master's Tools Will Never Dismantle the Master's House," in *Sister Outsider: Essays and Speeches* (Berkeley, CA: Crossing Press, 2007), 110-114.

at the expense of those who were not responsible for the failure. In this way, such failures, even if necessary for the possibility of success in some cases, reinforce the mechanisms of oppression that are partially responsible for the failures in the first place, specifically because the ramifications of the failure are not the same for all participants.⁶²

Lastly, this example highlights not only why the idealization of agent-neutrality is not viable for any model of shared intentions that aims to cover a range of cases – both simple and complex – but it also shows that theorists can run into similar problems even when including agential identity and social categories. While both race and gender play a significant role in this case, a more thorough analysis requires simultaneous attendance to issues of class and the role wealth plays in access to and standing within academic institutions (particularly white wealth), the particular relations of power that occur within academic institutions and their relationship to non-institutionalized forms of intellectual life and knowledge production, and the larger demand that those formerly excluded from particular conceptions of community would benefit from assimilation into those communities. All of these features determine the terms of success or failure in this case. One value of intersectionality as methodological and epistemological project is its ability to embrace the complexity of social identity, knowledge, and social categorization without oversimplifying while also being able to compare the complex interplay between particular contexts and globalized mechanisms of power.

Case 2: Opposing the _____ Together

On June 18, 2019, Representative Alexandria Ocasio-Cortez, of New York tweeted:

This administration has established concentration camps on the southern border of the United States for immigrants, where they are being brutalized with

⁶² This connects back to the ‘drop-in, drop-out’ participant whose presence might increase instances of failure but whose capacity to ‘drop-out’ easily protects them from the ramifications of that failure.

dehumanizing conditions and dying. This is not hyperbole. It is the conclusion of expert analysis.⁶³

Ocasio-Cortez's tweet was met with substantial pushback for its use of term 'concentration camps' as a replacement for 'detention facilities.' Ocasio-Cortez was accused of inappropriate hyperbole (despite arguing that her use was consistent with that of expert definitions) and anti-Semitism by Representative Liz Cheney, among others, who claimed that the use was demeaning to Jewish Holocaust victims.⁶⁴ Defending Ocasio-Cortez's use of the term in an article titled, "The Unimaginable Reality of American Concentration Camps," Masha Gessen points out that:

Like many arguments, the fight over the term "concentration camp" is mostly an argument about something entirely different. It is not about terminology. Almost refreshingly, it is not an argument about facts. This argument is about imagination, and it may be a deeper, more important conversation than it seems.⁶⁵

Gessen's point is Ocasio-Cortez's use of the term is not justified in terms of a set of necessary and sufficient conditions that allow us to distinguish between 'detention facility' and 'concentration camp,' but rather "about how we how we perceive history, ourselves, and ourselves in history."⁶⁶ Large-scale atrocities like the Holocaust become unimaginable, because in painting the action and its actors as monstrous, we effectively remove them from human history, making comparisons within human history impossible:

In crafting the story of something that should never have been allowed to happen, we forge the story of something that couldn't possibly have happened. Or, to use a

⁶³Alexandria Ocasio-Cortez, "This administration has established concentration camps on the southern border of the United States for immigrants, where they are being brutalized with dehumanizing conditions and dying. This is not hyperbole. It is the conclusion of expert analysis," June 18 2019, 8:03 a.m., <https://twitter.com/AOC/status/1140968240073662466>

⁶⁴Liz Cheney, "Please @AOC do us all a favor and spend just a few minutes learning some actual history. 6 million Jews were exterminated in the Holocaust. You demean their memory and disgrace yourself with comments like this," June 18, 2019, 9:25 a.m., https://twitter.com/Liz_Cheney/status/1140988893627478018.

⁶⁵ Masha Gessen, "The Unimaginable Reality of American Concentration Camps," *The New Yorker*, June 21, 2019, <https://www.newyorker.com/news/our-columnists/the-unimaginable-reality-of-american-concentration-camps>.

⁶⁶ Gessen, "The Unimaginable Reality of American Concentration Camps."

phrase only slightly out of context, something that can't happen here. A logical fallacy becomes inevitable. If this can't happen, then the thing that *is* happening is not it. What we see in real life, or at least on television, can't possibly be the same monstrous phenomenon that we have collectively decided is unimaginable.⁶⁷

Thus, disagreement about the use of the term cannot be recoiled by merely appealing to the facts of the matter because its use serves a different purpose: to change how we think about and thus respond to the policies and actions of the United States Immigration and Customs Enforcement agency (also known as I.C.E.).

In this way, it affects how people perceive what is happening *and* how people act in relation to it. Ocasio-Cortez's use of the term is akin to similar contemporary appeals to linguistic change, such as feminist calls to replace phrases like 'unwanted/non-consensual sex' to 'rape.' In these cases, the goal is to contest and alter the current conceptual schema in order to shift a number of relevant items. For instance: (1) changing our moral and legal language changes how we determine responsibility, blame, and punishment; (2) changing our moral and political language shapes how victims and survivors conceptualize what has happened to them, and by extension affects processes of retributive, reparative, restorative, and/or transitional justice; (3) how the events are discussed in the media and documented for historical record; and (4) how we ought to respond to the events as they are occurring.⁶⁸ This is why, for example, the

⁶⁷ Gessen, "The Unimaginable Reality of American Concentration Camps."

⁶⁸ See: Bill E. Lawson, "Moral Discourse and Slavery," in *Between Slavery and Freedom: Philosophy and American Slavery*, eds. Howard McGary and Bill E. Lawson (Bloomington: Indiana University Press, 1992), 71-89. Lawson argues that the lack of a word to capture both the institution and aftermath of slavery in America alters our moral discourse in terms of how the ongoing impact is understood, how the demands of justice are discussed, and how harm and related forms of damage are conceptualized. For example, the term 'Holocaust' radically alters how perpetrators, victims, and bystanders understand Nazi Germany's genocide and how Americans talk about the demands of justice in this case.

difference between ‘mass murder’ and ‘genocide’ are relevant for considerations of shared or collective responsibility.⁶⁹ It changes the nature and meaning of the action.

Consider, now, the differences between two statements that both refer to the same physical building: (1) “We intend to oppose that concentration camp,” versus, (2) “We intend to oppose that detention facility.” If two individuals, who are trying to undertake a joint action in relation to the events taking place within that physical building, each individually hold (1) and (2) respectively, then there is a question of whether they, in fact, have a shared intention.

Following Gessen’s analysis, there is reason to think that they do not, because the difference in terminology changes the nature and meaning of the action. For example, it changes the urgency with which we should act and changes the landscape of what kinds of actions we are justified in taking. The term ‘concentration camp’ communicates a need for urgent response to the situation, which shifts the process of deliberation in developing a joint action. For Virginia Held to argue that people can share responsibility for failing determine a decision-procedure quickly enough in order to swiftly decide on and carry out a joint attempt to save a drowning child, her argument must rest on an appeal to urgency.⁷⁰ The strangers on the beach do not have limitless time to decide how to save the child; they must act and they must act now. For joint actions, the urgency also requires us to quickly determine who has the relevant expertise, whose knowledge and/or which standpoints are most useful, and why. It is not merely sufficient, as both Kutz and Bratman’s models dictate, that participants meet common knowledge or mutual openness conditions and intermesh their sub-plans accordingly. In order for joint action to be successful

⁶⁹ See: Tracy Isaacs, *Moral Responsibility in Collective Contexts* (New York: Oxford University Press, 2011), 12, 123-127. Tracy Isaacs also notes that this affects individual responsibility as well. For example, there is a difference between someone murdering a Jewish person and someone murdering a Jewish during a genocide that targets Jewish people.

⁷⁰ Virginia Held, “Can a Random Collection of Individuals be Morally Responsible?” *Journal of Philosophy* 67, no. 14 (1970): 471-481.

under conditions of urgency, there is a need to ensure that there is agreement about the nature and meaning of the action, and further coordination once the intention is in place that appeals to the relevant knowledge or expertise, which may not be equally shared by all participants.

Moreover, the term also communicates a need for a different kind of response, a different kind of action, for it appeals to historical knowledge about what is most effective and offers a different moral framework for what kinds of actions might be justified. Historically, to oppose a concentration camp is to liberate it and to resist political conditions that make it possible by any means possible. Opposing a detention center, on the other hand, does not necessarily carry those connotations (even if it ought to); to oppose is to, perhaps, appeal to law, to change the policies, to engage in non-violent protests.

In this way, there is an open question about the role of language in what it means to share an intention. If the two statements do not represent the same intentional content, then one must ask how similar intentional states must be in order to possibly be shared. Both Bratman's and Gilbert's models do not provide an answer to this problem, but part of their idealization of the simple case often refers to a simple action in which the agents already have a sufficiently shared understanding of what the action is, such that minor linguistic differences in intentional content might not cause any issue, if they are present at all. For different reasons, Kutz's model allows for differences in intentional content – for example, between the intended goal and the intention to do one's part – and his reliance on a shared conceptual background required for mutual understanding will also likely allow for minor linguistic differences.⁷¹ However, one might

⁷¹ However, one might argue that the intention to do one's part requires some level of conceptual agreement with the intended goal.

argue, contra Kutz, that the intention to do one's part requires some level of conceptual agreement with the intended goal.

However, the difference between 'concentration camps' and 'detention facilities' highlights that being able to pick out the same intended referent ("That building was those people in it") is not sufficient for a shared intention.⁷² This problem is further compounded by the fact that in such cases there will likely be disagreement amongst potential participants as to whether the terms are sufficiently similar. For instance, one might say, "I don't care what we call them. I just care that we shut them down," whereas another might then point out that how the participants refer to them impacts how they go about shutting them down. More generally then, the issue is that both linguistic differences and differences in conceptual schemas must be sufficiently understood and dealt before a shared intention is possible. However, because Bratman, Gilbert, and Kutz all use idealizations that assume that there is a shared conceptual background between agents, or build their models from cases in which such failures would not be possible (or would simply paint one of the participants as irrational or lacking in some agential capacity). This cannot be used in a model if the goal is to capture a wide range of different cases (i.e., Gilbert and Kutz) or if the goal is to be able to further extend the basic model (i.e., Bratman).

However, even if one claims that the intentional content must match perfectly or nearly perfectly, this does not resolve the problem for Gilbert. To use her example of walking together, suppose two people both intend to go for a walk together, and in fulfilling all of Gilbert's

⁷² This only further highlights why the idealization of agent-neutrality is not viable because it opens up further questions about terminological differences in relation to social categories and identities both in cases where one term might function as a slur (e.g., 'Jew') or has fallen into disfavor (e.g., 'homosexual' v. 'gay'), and when the term does not match the person's identity (e.g., the person in the example does not identify as a woman, even if others perceive them to be). Certainly, in the case of anti-oppression resistance, such concerns around terminological differences become even more of a pressing issue.

criteria, both people hold that “We intend to go for a walk together.” Now suppose person A starts walking at 2.0 mph and person B starts walking at 4.0 mph. Unlike Bratman and Kutz, Gilbert does not have a success condition about inter-meshing of sub-plans or strategic responsiveness, so we can successfully share the intention to walk together, even if we do not coordinate the pace. The problem is thus how to determine whether the joint commitment has been broken and who broke the commitment. At some point, person B will be so far ahead of person A that it might be odd past a certain point for either of them to have a sense that they are walking together, but it is not clear at what point the joint commitment was violated (or if it was) nor who broke it.

Returning to the case, the pushback against Ocasio-Cortez’s use of the term also is a debate about who gets to determine when the term is being appropriately applied. That is, there is a question of when a conceptual shift is required, who makes that determination, and whether (and, if so, why) are some individuals or groups are in a better epistemic position than others. This connects back to the question of who has the relevant expertise, whose knowledge is most useful, and who gets to speak on behalf of whom. Liz Cheney’s tweet intended to speak on behalf of Jewish people but was swiftly met by opposing by members of that very group, arguing that not only is the term correct in this case, but that the only reason the slogan “Never Again” exists is to be able to correctly identify escalation towards genocide (or existing genocide) and to communicate a duty to oppose the mechanisms of genocide, i.e., concentration camps.⁷³ Of

⁷³ See, e.g., “We Are Not Fucking Around,” *Never Again Action*, accessed on February 24, 2020, <https://www.neveragainaction.com>. As the section reads: “We are not going to sit at home while people are dying and families are being rounded up. We are not going to wait and see what happens. We are going to keep building this movement of Jews and allies until we’re strong enough to make it impossible for ICE to function and for the camps to stay open. We owe it to the immigrants we’re in this fight with, and we owe it to our ancestors who gave us a sacred mission: Never Again for Anyone.” Although, one might rightly refer back to the first case and question whether this claim (and the actions they have followed) also reflects shared intentions in the case of solidarity or failed in a similar capacity.

course, there is not agreement between all those who are Jewish, and Jewish people do not have the sole right to determine when its use is appropriate. However, the fact that many Jewish people, particularly scholars of the Holocaust, survivors, and their descendants, have defended Ocasio-Cortez should be one reason for taking her use seriously, because Jewish communities have a relevant expertise due both a set of historical experiences and deep cultural practices on reflection of those experiences, which can constitute a standpoint.⁷⁴ Of course, the question of who makes the determination most pressingly includes those who experienced these American camps directly as prisoners, indirectly in their relationships to the imprisoned, or as a potential target (regardless of citizenship status) of American white supremacy and xenophobia. That is, the determination of relevant expertise and knowledge tracks social-structural positioning in some capacity.

The issue here is not merely about the use of the term, but how its use connects to other conceptual schemas about forms of oppression and forms of resistance. There is a reason that people often are tepid about any analogies to Nazism. If the analogy is appropriate, then it radically shifts the framework for joint action, the appeal to particular forms of expertise and knowledge, our framework for moral, political, and legal responsibility, and how people understand themselves within ongoing structures of oppression. It changes how we think about acting together and thus it changes how we actually act together. Conceptual schemas impact our agency by shaping what we take agency to be, by shaping people's self-understanding, sense of group membership, and sense of others through processes of social categorization, and by shaping the nature and meaning of our actions, either individual or joint.

⁷⁴ Because it constitutes a standpoint, it does not matter that some Jewish people oppose the use of this term, because a standpoint is a claim about group, not individual, knowledge. In fact, standpoints in themselves might constitute a form of idealization.

Section IV: Concluding Remarks

These examples thus highlight the myriad ways in which the current idealizations used in Bratman's, Gilbert's, and Kutz's models for shared intention produce inferior models, given that their purpose to capture a wide range of cases and/or capture small-scale cases under the assumption one can likely scale-up the model. The idealization of agent-neutrality cannot account for both the construction of human difference and the impact that agents' social-structural positionings has on joint actions. An oppression-excluding, idealized social ontology cannot address how previous and ongoing forms of reality mismanagement shape our understanding of shared agency, nor can it help us manage issues that arise in trying to share agency with others against a backdrop of structural oppression. The ways such models idealize epistemology and epistemic practices cannot accommodate clashes in conceptual schemas between agents, nor can it account for how conceptual schemas often track social dimensions in the creation of knowledge. The problem is not merely that we end up with inferior or poor models, but rather that these models aim to manage the social reality of shared agency without addressing whether we ought to manage them in this way. If processes of social construction and structural oppression play constitutive roles in shaping shared agency and joint action, then the ethical question is also an ideological concern. If our models are not able to include the concerns raised by oppression theorists, then they do not remain neutral. That is, these models will likely rely on ethically or politically concerning conceptual schemas under the guise of "neutrality" or "universality." In doing so, these models become useless for understanding joint actions that aim to resist structural oppression (and why these actions succeed or fail), and risk reinforcing the very conceptual schemas central to the mechanisms of structural oppression.

Chapter V

Model Pluralism and The Ethics of Social Reality Management

Section I. Introduction

The study of shared agency within the analytic literature on shared intentionality has largely followed an ideal methodology. In developing a model for what it means for two or more people to successfully share agency, theorists tend to rely on three unjustified forms of idealizations. First, the idealization of agent neutrality posits an asocial picture of a subject in which an agent's primary feature is rationality. Second, an idealized social ontology that excludes ontological injustices born from ongoing, dynamic, and necessarily co-forming systems of power and oppression. Third, an idealized epistemology that offers an asocial and apolitical approach to the production of knowledge and in which agents are assumed to operate off of significantly shared ideological and conceptual frameworks, unaffected by structural forms of ignorance. My argument thus far has focused on the methodological and ideological contours of these highly idealized and general models for shared agency. Drawing on Hans Vaihinger, I have argued that the problem with an ideal methodology in this case is two-fold. First, because models for shared agency are normative, the application of an ideal model treats the real world as a deviation to be fixed in accordance with the model, not a signal that the model itself might be inadequate. Second, an ideal methodology relies on bad, or contextually unjustified, idealizations, resulting in models that not only fail to increase our understanding of the ways in which shared agency arises within our social worlds, but also relies on parochial values and

metaphysical assumptions that, as Charles Mills argues,¹ reflect the interests and worldviews of those who largely benefit from overlapping systems of power and privilege.

Oppression theorists, such as myself, have historically been critical of the use of idealizations because idealizations often exclude the concrete, complex realities of structural oppression and thus assume that any models relying on idealizations cannot help us ameliorate oppression. Rather than merely shifting to a non-ideal methodology in which idealizations of any sort are eschewed, I have shown that idealizations are necessary untruths that allow their users to grapple with the complexity of dynamic systems of power. Ideal theorists idealize away from complexity by removing the existence of dynamic social categories as well as local and global formations of structural oppression, but oppression theorists use idealizations instead to capture such complexities by including contextually limited idealizations that track the particular formations of social categories and the mechanisms of power across different but highly interconnected social-ontological landscapes. Understanding how oppression theorists use idealizations thus highlights why the three prominent models of shared intentions rely on bad idealizations, which results in inferior models. These models cannot explain the terms of failure or success for shared intention and joint actions in the case of anti-oppression resistance, nor can they capture the potential variety of forms of shared agency.

Thus, my argument lays out three overlapping concerns about these idealized models: methodological, ethical, and ideological. First, my methodological concern is that the aforementioned idealizations in models of shared agency exclude the role of structural oppression, resulting in inferior models. That is, they offer an inferior explanation of shared agency and joint action compared to a model whose idealizations incorporate the role of

¹ Charles W. Mills, “‘Ideal Theory’ as Ideology,” *Hypatia* 20, no. 3 (Summer 2005): 165-84.

structural oppression. Moreover, in virtue of their normative elements, these models offer an account of what joint actions are and how agents ought to act within them. The problem is that if these models offer an inferior explanation, they not only fail to capture instances of the phenomenon that do not match their models, but also limit the possibilities in terms of what constitutes shared agency and joint actions.

Second, my ethical concern is: given that shared agency and joint actions are their own form of reality management – that is, the management of our social reality and ourselves within it – then the methodological issue is also an ethical one. My ethical concern has two dimensions. In the first case, there is a question of whether the purpose of such models is ethical and whether the conceptual schemas on which that models relies are themselves ethically concerning. These two questions are not separate and answering them depends on how the two relate within a particular context. That is, whether the purpose is ethical depends on the scope and context of application, whether the context of application is within the conceptual schema on which it depends, and whether there are ethical concerns about that conceptual schema (or features thereof), both internally and externally to the schema. In the second case, the question is not only about how we ought to manage shared agency and joint actions, but who is justified in this management and why. Answering this question, again, depends on the content and scope of the management. Lastly, my ideological concern is: when models for shared agency use idealizations that actively exclude the impact of structural oppression, they risk adopting the outcomes of previous attempts at reality management, which have played a vital role in the creation and maintenance of oppression and that reinforce those very forms of oppression. This makes it impossible, among other things, to use these models for shared agency and joint action to resist structural oppression. Moreover, it means that using such models as a form of reality

management is likely to reinforce structural oppression by reinforcing other ongoing forms of reality mismanagement (i.e., unjustified management).

In this final chapter, my goal is to bring together my methodological, ethical, and ideological concerns. I argue that in the case of shared agency, there is a methodological need for a multiplicity of models or, more specifically, an *open-ended model pluralism*. I analyze the ethical and ideological dimensions of this methodological need by arguing that a *constrained model pluralism* constitutes a tyrannical form of second-order social reality management. Drawing on analysis of colonial forms of feminism, I show how constrained model pluralism (and the idealizations often used with it) results in a tyrannical form of social reality management, even in cases where those, like the conference organizers of chapter IV, who attempt to engage in shared agency in the case of transnational struggles, can still engage in unjustified forms of management. Again, this serves to highlight that we not only need to examine whether a model's purpose seems ethical on the face of it, but also examine what the actual impact of the model is for those being managed by the model. Lastly, drawing on later Wittgenstein, I argue that open-ended model pluralism treats forms of shared agency as different language games. If our goal is to develop models for massively shared agency when needed, then it is the friction between different models for shared agency that allows for the development of such models.

Section II. Open-Ended and Constrained Model Pluralism

Open-ended model pluralism, broadly, is the claim that theorists of shared agency should aim toward a multiplicity of models, particularly focusing on more localized or contextually specific forms of shared agency. If the goal of research on shared agency is to both understand and manage this phenomenon, then descriptive philosophical accounts require a more

interdisciplinary approach in order to genuinely engage with the extraordinarily wide range of forms of shared agency. Of course, understanding and management are intertwined. There is no actual distinction between descriptive and normative models for shared agency, as such models necessarily reflect particular values and ontological assumptions. Thus, my claim is not that philosophers do sociology, for example, but that in developing models, we need to start from a particular context and doing so requires engagement with the particular socio-ontological landscapes and the histories that have shaped them. This includes critical reflection on how our own specific values and ontological assumptions appear in our models in order to, at minimum, have a clearer picture of the purpose of the model and its potential limitations. Open-ended pluralism also demands such investigations are not pre-constrained by non-universal standards of reasonableness (akin to the constraint of reasonable value pluralism seen in John Rawls' *Political Liberalism*),² as this genuinely limits not only our understanding of what shared agency is but, more pressingly, what it could be. In this way, open-ended pluralism ought not seek to merely to understand and manage forms of shared agency that currently exist, but to use this pluralism to figure out what possibilities for shared agency there might be. That is, shared agency research should not restrict itself to what is by incorporating questions about what shared agency could be. In this case, I argue that open-ended model pluralism provides a stronger basis for deriving more general models in cases where forms of massively shared agency are needed, ensuring, however, that these general models are still contextually tailored to the particular needs for mass action.

In contrast, the critical part of my project has focused on *constrained model pluralism*. This constrained pluralism is, broadly, the assumption that there is only need for a limited

² John Rawls, *Political Liberalism*, expanded ed. (New York: Columbia University Press, 2005).

number of highly general or universal models, in which the nature of the limitation is predetermined by particular standards of reasonableness and/or a picture of agency, i.e., agency monism. Moreover, constrained pluralism does not investigate the future possibilities of shared agency. This is directly tied to the use of an ideal methodology. Idealizing away from the particulars of context is to take an atemporal approach to shifts in the varieties of shared agency over time. An ideal methodology tends to rely on the assumption that it will produce the best (i.e., most ideal) models, and thus cannot ask questions about betterment without appeal to a limited, parochial ideal, nor ask questions about possibilities that in no way refer to this static ideal. Lastly, constrained pluralism fails to treat the process of theorizing and developing models of shared agency as something we necessarily do together (often by examining the similarities and differences between more localized models). In this way, the claim that community is a site for creative social action and that social action is a form of theorizing in itself highlights how theorizing shared agency always happens through particular communities, that theorizing shared agency is already something we do with others – whether it be in more explicit or implicit ways.³ In this way, theorizing shared agency – the development of a particular model – cannot be something we do alone through an idealized methodological procedure that removes the complex and dynamic features of our sociality.

³ Collins, *Intersectionality as Critical Social Theory*, 157-188. In her discussion of community as the site of creative social action, Collins is pulling directly from the tradition of American pragmatism: “Because experiencing the social world is always subject to interpretation and reinterpretation, identities are social phenomena and as such are never finished. Rather, experiences and the identities they engender are always in the making; that is, people shape their social worlds through the actions they take as well as the experiences that their actions engender” (173). Collins notes that the experimentalism of pragmatism highlights how our actions take on meaning contextually but also that our concepts that drive our actions take on meaning through actions themselves, allowing us to test and retest our experiential understandings and concepts through acting together, particularly within the context of community.

Section III. Constrained Model Pluralism

My claim is now that the three prominent models analyzed represent a form of constrained model pluralism in virtue of their ideal methodology and bad idealizations. However, one might challenge my argument here that given my claim about open-ended model pluralism, which following Vaihinger's approach will necessarily include incompatible models – Bratman's, Gilbert's, and Kutz's models ought to be included within an open-ended pluralism, not rejected outright.

My response to this challenge is two-fold. First, one central goal of my critique has been to raise concerns about the questionable relationship between the purpose of these models and the particular idealizations used within them. The endpoint of this critique is not, however, a full-scale rejection of these models. Nevertheless, the inclusion of these models would likely require revised versions. Full-scale rejection would make the very move I am critiquing, namely that there is one set (or a limited set) of values that ought to guide research on this topic. The goal of this methodological and ideological critique, rather, is to assess what values inform such models and determine the appropriateness of such values in relation to the purpose of that model *as well as* assess whether those values are, for further reasons, objectionable in their own right. One might develop a very effective shared agency model for manufacturing a genocide, but we would also have independent reasons for claiming that this purpose is morally objectionable.

Returning to the particulars of the three accounts, let's take Gilbert's joint commitment, for example. If the purpose of Gilbert's model is to capture the basic structure of all possible instances of shared intentions, and the model relies on the idealization of agent-neutrality, then it seems like such a model will not only be inferior (and in many cases useless), but also the value of neutrality risks universally imposing a particular conceptual schema for agency. It thus

excludes different conceptual schemas for agency *and* the imposition itself risks radically altering what it is imposed upon, providing a form of de facto evidence for the model's usefulness. Notice that even if there were no risk of imposition or one were to argue such an imposition is justified, the model would still fail to provide an enhanced understanding of how shared intentions and joint actions work outside of this particular conceptual schema, nor would it be able to explain how agents across different schemas seem to successfully share intentions. However, with alteration, Gilbert's criteria of personal readiness might be useful for describing low-level forms of shared agency across difference, even when there is a lack of shared language. For example, a modified version of personal readiness might explain how two people without a shared language could jointly commit to moving a heavy piano together down a flight of stairs, and why one party might be justified in rebuke if the other simply dropped the piano halfway through and walked away.

While Kutz's model for participatory intentions also purports to have universal application, his reductionist form of participatory intentions might better capture spontaneous formations of joint actions in which there is a clear first actor who acts without knowing whether others will also start to participate. Neither Bratman nor Gilbert's models can capture this particular kind of spontaneous emergence. In comparison, Bratman's model might prove valuable for capturing how small-scale forms of shared planning agency works between adult agents who have: (1) a significantly shared social background, and (2) a relatively equal social-structural positioning within a particular context.

In this way, each of these models seems to capture different elements of the many ways in which shared agency occurs but their usefulness in this regard requires a revision and contextualized narrowing of their purpose. Gilbert and Kutz, for instance, would need to give up

on an universalizable model. Bratman, on the other hand, might need to narrow the application of his model to agents who already have sufficiently shared norms. For this reason, and in line with open-ended model pluralism, my claim is not that we ought to reject these models outright, but rather, the scope of such models needs to be significantly narrowed, including attention to how their parochial assumptions about agency might only apply to those who share those assumptions – particularly insofar as one’s understanding of agency impacts one’s sense of agency, either individually or shared. Each does seem to capture valuable features of different forms of shared agency, even if those features are not, nor should be, universalizable.

That said, this response only addresses the explanatory purpose of these models, not their purpose of management, and this complicates the optimism in the first part of my answer. The problem is not the fact of second-order social reality management in itself, as surely this can be a justified practice (e.g., the goal of social justice is one such version of this practice). Rather, the problem is about the scope and ethical nature of this management. More specifically, the problem is that in attempting to manage features of social reality, e.g., attempting to resist a particular form of injustice, it can be relatively easy to uncritically adopt the unjustified or morally unacceptable ontological outcomes of previous management attempts (regardless of whether those attempts were justified in their own right). I take this to be a particularly pressing concern in the case of structural forms of oppression.⁴

⁴ This is not to say that structural oppression is a social form of reality management in itself because structural phenomena are not reducible to the intentions of individuals or institutions. However, it is to say that the mechanisms of structural oppression are only made possible by forms of reality mismanagement, which is why even forms of justified reality management within structural oppression can produce morally unacceptable outcomes. See: Iris Marion Young, *Responsibility for Justice* (Oxford: Oxford University Press, 2011).

Section IV. Management and Justification

Shared intentions and joint actions are, in their own right, a form of social reality management on the part of the participants. Protesting together is about managing our political reality, developing feminist theory together is about managing our social reality by managing the conceptual schemas that shape our social world. In philosophy, the practice of conceptual engineering is not merely about asking what social reality is, but instead asking how it should be or, differently, what we want it to be.⁵ Helpful to my purpose here, Sally Haslanger argues that in asking what race or gender *really are* misses the point: “There are many different types of human bodies; it is not the case that there is a unique “right” way of classifying them, though certain classifications will be more useful for some purposes than others.”⁶ That is, the purpose and function of these classifications varies depending on the context. In the case of unjust socio-ontological landscapes, we are called upon to revise or discard classifications that have ceased to be valuable (if only for resistance to the classification), or actively respond to forms of classification useful only for the purposes of domination and/or retaining unjustified access to power. Of course, who gets to make those determinations, in what contexts, and why are far more complicated questions.

What Haslanger’s piece also illuminates is that this practice of conceptual engineering is already a shared social practice – one akin, perhaps, to forms of group knowledge production. In this way, conceptual engineering is a form of shared social reality management. It is something

⁵ See: Herman Cappelen, “A Guided Tour of Conceptual Engineering and Conceptual Ethics,” in *Conceptual Ethics and Conceptual Engineering*, eds. Herman Cappelen, David Plunkett, and Alexis Burgess (Oxford: Oxford University Press, 2018), 3. Cappelen defines conceptual engineering broadly as: “(i) The assessment of representational devices, (ii) reflections on and proposal for how to improve representational devices, and (iii) efforts to implement the proposed improvements.”

⁶ Sally Haslanger, *Resisting Reality: Social Construction and Social Critique* (Oxford: Oxford University Press, 2012), 246.

we do together. Herman Cappelen notes how this might work in the case of social ontology: “If lexical items play some role in the creation of social facts, then changing the meaning of a lexical item might contribute to a change in social reality.”⁷ While I do not take it that lexical changes necessarily beget socio-ontological changes, I do think that Cappelen underestimates the potential power of lexical changes in this case. Take, for example, the function of marriage as a legally regulated form of shared agency in the United States. The legal concept of ‘marriage’ has been revised to include a broader range of marital arrangements, but the terms of this form of shared agency have also changed. The introduction of the concept of ‘marital rape’ and the subsequent outlawing thereof, altered the legal terms of this joint undertaking. By qualifying the term ‘rape’ with ‘marital,’ those responsible for the development and use of the term were able to make explicit a long-standing, patriarchal assumption that the legal bond of marriage precluded the possibility of rape therein (which tracked the additional assumption that heterosexual marriage was hierarchically arranged to the disfavor of the woman). Conceptual engineering in the case of marriage shows that shifts in language do, in fact, alter our social reality, at least within particular contexts. This is not to say that the changes necessarily reduced instances of sexual assault (that is an empirical question),⁸ but rather shifting the legal nature of arrangement made explicit, in a beneficial way, the unjustified assumptions about the precise terms of this contractually shared agency.

I take the case of conceptual engineering to show that shared intentions and joint actions are not only a form of social reality management, but also the site at which we can critique,

⁷ Herman Cappelen, “Conceptual Engineering: The Master Argument,” in *Conceptual Ethics and Conceptual Engineering*, eds. Herman Cappelen, David Plunkett, and Alexis Burgess (Oxford: Oxford University Press, 2018), 23.

⁸ The shift did, however, provide new legal ground to address sexual assault within legally constituted marriage, even if the legal system within the United States fails, for independent and structural reasons, to offer justice for survivors of sexual assault.

resist, or otherwise respond to past or ongoing forms of management that we take to be ethically and/or ideologically concerning. Given the arguments about methodology and epistemology in Chapter IV, what ‘we’ take to be ethically concerning must track questions about power-infused practices of knowledge production and structural forms of ignorance. There are many ways in which forms of shared agency ostensibly aimed at managing forms of injustice are: (1) actually unjust forms of management, e.g., imperialism, or (2) the replication of unjust practices in the attempt to respond to genuine forms of injustice.

This is why the purpose of social reality management is a pressing concern for how theorists develop models of shared agency. These models are a second-order form of social reality management; that is, they manage the terms of this management. This is where my methodological, ethical, and ideological concerns dovetail. Bratman, Gilbert, and Kutz all rely on idealizations that do not provide a better understanding of how mechanisms of structural oppression shape the manifestations of and possibilities for shared agency. At the same time, these models are a form of second-order management for our non-ideal, complex, and power-infused forms of shared agency. All three rely on an appeal to neutrality, but as argued in previous chapters, this is not genuine neutrality but a way to mask the values (including the value of neutrality) and conceptual schemas that inform what we take agency to be and the ways in which it functions. This is why oppression theorists have (albeit mistakenly) sought to rid theory of idealizations entirely. The worry is that such idealizations necessarily ignore ethically, socially, and politically pressing concerns related to structural oppression, specifically insofar as the idealizations reflect those who largely benefit from the mechanisms of oppression. Thus, shared agency theorists need to critically examine both whether the particular idealization

ensures that the model more successfully fulfills its purpose *and* whether the purpose itself is ethically or politically justified.

There are two kinds of justification in this case. First, the purpose of a model is justified in terms of its context. For instance, if the purpose of a model is to understand how forms of shared agency work within a cultural milieu that offers a hyper-individualistic and self-interested picture of agency, then idealizing in virtue of how agency is understood within this context will likely provide a better explanatory account of how successfully shared agency occurs between those whose subjective sense of agency reflects this picture. However, because the purpose is about management and not merely explanatory power, one might ethically or politically assess the purpose of a model that aims to enhance such forms of shared agency without reflection upon whether we ought to reinforce this picture of agency in the first place. The question is then whether the goal of this management is to maintain or change the forms of shared agency in question, and on what epistemic grounds we start to question whether certain forms of shared agency (in virtue of the broader picture of agency they presuppose) ought to be changed and why. One might argue, for example, that it is valuable to understand how shared agency works when agents understand themselves to be essentially self-interested but nevertheless ethically or politically concerning to reinforce this kind of conception of agency.

The second case of justification asks whether the purpose itself is justified in virtue of who is engaged in this form of second-order social reality management. Thus, the ethics of social reality management is about who gets to manage what and why. It's worth returning here to Iris Marion Young's distinction between traditional and modern forms of oppression.⁹ On a traditional model, oppression occurs when one person (i.e., a tyrant) or a group of persons

⁹ Iris Marion Young, "Faces of Oppression," in *Justice and the Politics of Difference* (Princeton, NJ: Princeton University Press, 1990, 41-42.

exercises intentional and unjustified control over others. The creation and maintenance of this control constitutes a form of reality management. This is often done through a combination of direct and indirect forms of violence (e.g., assault or manufactured starvation) as well as shaping the kind of conceptual schemas that tend to ensure sufficient compliance and induce fear in those being controlled. Structural oppression, on the other hand, occurs when various forms of unjustified reality management interact, modifying each other in those interactions, and often relationally increasing their effects. However, the mechanisms of structural oppression also interact with and alter otherwise justified forms of management.¹⁰ This is why the ideological critique is important. The complex mechanisms of structural oppression make it harder to discern which and to what extent various forms of management are replicating forms of domination and social inequality. By extension, this makes it harder to discern to what extent forms of unjustified management (and their outcomes) are taken up within the purpose of a particular model as well as within the model itself.

While this problem is not unique to shared agency, it does not make the problem any less pressing. The problem of discernment circles back to the discussion of feminist philosophy of science and feminist epistemology in Chapter VI. The ethics of social reality management concerns both researcher-researched and knower-knowledge dynamics, and calls such theorists to critically reflect upon the relationship between the managers and the managed. On a traditional model of oppression, the problem is with a single tyrant or a group of rulers, such forms of tyrannical reality management are often easier to notice and, given that tyranny is not

¹⁰ This is why the question of responsibility or duties to resist structural oppression can be complicated. At the University of Memphis' Philosophy Graduate Student Association's 2015 conference, "Structural Oppression and Environmental Ethics," Kyle Whyte aptly referred to structural oppression as a 'whack-a-mole' problem. In the game 'Whack-A-Mole,' the attempt to push down one popped-up mole begets the popping up of another. It's a no-win situation. That is, the attempt to solve the problem in some cases always results in the reinforcement of the problem elsewhere, or the attempt to solve one problem necessarily begets the reinforcement of another.

justified, easier to denounce as a form of mismanagement. When oppression is structural, however, intentional forms of reality management are decentralized and their effects caught up within large structural mechanisms. It makes it harder to see how ostensibly descriptive models of social reality are already normative, and thus harder to determine and denounce forms of mismanagement, especially when certain groups benefit from ongoing mismanagement.

This returns to the earlier claim that theorizing shared agency and forms of social reality management are practices that people engage in together in contextually specific ways. However, when Haslanger writes in theorizing about race and gender that “we should begin by asking (both in the theoretical and political sense) what, if anything, we want them to be,”¹¹ this ignores a more pressing question: to whom is the ‘we’ referring? That is, who gets to decide what gender or race should be, in what context, and for what purpose? Answering such questions requires a methodological approach infused with a non-ideal epistemology from the start, tracking claims about epistemic advantage, the relationship between power and ignorance, as well as the epistemic practices that arise from critical reflection on the methodological process itself and its outcomes. Certainly, most people do not think that a single person has the right to determine what race or gender should be, but it also does not follow that everyone is equally justified in trying to make that determination. That is, it seems that at the minimum, the justification for social reality management is dependent on context (i.e., what is the purpose of the model and its scope of application?) and requires appealing to concerns regarding epistemic advantage and ignorance.

The examples in Chapter IV highlight the ways in which the modeling of shared agency connects with ethics of this management. The all-white organizers for the *Signs* conference

¹¹ Haslanger, *Resisting Reality*, 246.

provided a model for ‘community’ that did not merely reflect their particular experiences but was also intended to reflect a more universal experience among women. Even in reflecting upon their failures, the organizers argued that they had been successful after all in creating the very community they thought necessary for reflection on the topic of ‘community.’ In this way, the organizers continued to define ‘community’ on their terms by integrating their failures into an ultimate success for everyone involved and therefore still imposing their model in documenting the conference by showing how, in the end, their model of community was successful. There is never a question of whether they were in the position (or at least alone in the position) to decide the success of the conference and what the success conditions were in the first place. In this case, the problem is not only that they excluded other participants from making this determination but also that they excluded invited participants who turned down the offer because of the organizers’ failures. By placing themselves at the center, by being the ones to extend the invitations, they already determined who had the chance to be included in *their* community in the first place and thus whose testimony was relevant to the question of success.

Similarly, the question of terminology in the case of American concentration camps is a question of who gets to manage the political reality of such places through the act of naming. Those who argue against the use of the term ‘concentration camps’ want to manage the reality of these places (and what that reality tells us about our larger social and political culture in the United States) differently than those arguing for the use of the term. The use of the term is not merely about managing the reality of the camps, but also about managing how people respond to them, what form those responses take, and how we document these camps and situate them within history. In this case, the question is who gets to decide on the terminology and why those

groups are justified in doing so; that is, why are they justified in that form of reality management in this particular case?

My discussion about the ethics of social reality management highlights why the purpose of a model for shared intention is not merely a question of usefulness, but also an ethical and ideological concern. There are certainly models for race and gender that are useful for creating and maintaining unjustified hierarchies *and*, in some cases, that forms of justice (e.g., transitional, rectificatory, reparatory) will need to refer to those models to ensure that such projects accurately track those harmed by the initial models (and thus the models can still be useful for this end even if that's not their intended purpose).¹² This is why I disagree with Appiah, who argues that we need not worry about talk of 'reality management,' that models built for unethical purposes will not be truly useful.¹³ A model for enacting a genocide is morally unjustified, but it would be a mistake to argue that the lack of justification renders such a pernicious model practically useless. Again, this requires critical attention to the dynamic between the manager and the managed because this relation can easily (and often does) become a dynamic between the dominators and the subordinated. Thus, shared agency theorists should be called upon to justify the purpose of their model in relation to a particular context and explain why they are justified in their attempt to manage some feature of social reality.

Lastly, this highlights why, compared to an ideal methodology, an intersectional approach offers a set of methodological and epistemological practices that get to the heart of concerns about shared agency, theorizing shared agency, and social reality management. In

¹² In this way, Haslanger's call for different models of race and gender is about the context in which they are used. For example, how someone identifies as Jewish shapes what kind of Jewish communities they seek out but whether someone who is ethnically Ashkenazi identifies as Jewish at all is irrelevant in certain medical contexts. Given that those who are Ashkenazi and biologically labeled 'female' are more likely to have *BRCA 1 and 2* gene mutations, a cultural or religious identity model for Jewishness, is not a useful model in this case.

¹³ Appiah, *As If*, 182n29.

particular, an intersectional approach offers three deeply valuable commitments for analyzing what it is to theorize shared agency, what it means to take up more particular purposes of social reality management (and why many can be unethical), and what this process might look like more particularly. First, an intersectional approach, like other forms of critical social theory, necessarily takes up social justice and social change as its purpose.¹⁴ However, in reflection on this purpose, Patricia Hill Collins highlights how social reality management devoid of commitments to justice results in national social engineering projects, like eugenics. Broadly, what shared agency theorists ought to take away here is that being engaged in any form of social reality management – be it first or second order – requires an ethically and politically grounded framework that intentionally limits the scope and potential forms of management. That is, it is not simply that particular forms of shared agency, such as those arising in cases of anti-oppression resistance, need to adopt social justice as their goal, but rather that social reality management is inherently political and as such is rife with the possibilities for larger-scale domination and violence (and, in fact, is often used toward that end). In this way, social and political justice is not a side-note to this methodological process. It is a necessary starting point for ensuring theorizing shared agency, as a form of second-order social reality management, is held accountable to concerns of power and domination at every stage in the process.

This commitment of purpose, in turn, tracks a commitment to both (1) reflexive accountability and (2) a dialectical and empirical approach (one that Collins’ compares directly to abductive analysis).¹⁵ Reflexive accountability broadly means that we are responsible for our methodological approaches, the ensuing claims to knowledge they generate, and the impact of

¹⁴ Collins, *Intersectionality as Critical Social Theory*, 21-53, 253-285.

¹⁵ Collins, *Intersectionality as Critical Social Theory*, 54-65, 147-157.

those claims on others. That is, responding to Haslanger, we are responsible for our claims about what race and gender are *as well as* our claims about what they ought to be. In this case, since developing a model for shared agency is a second-order form of social reality management, then we are accountable to the models we produce as theorists, their uses in terms of management, and the outcomes of such management. For example, white feminists are accountable to the ways in which our models of patriarchal oppression over-idealize the category of ‘woman’ in white-centric ways, but also for the ways in which, due to racial privilege, such theories were more broadly taken up as authoritative, as accurately reflecting the ontological structure of gendered oppression, and then for the impact on people’s actual lives when such theories were applied. Comparatively, an ideal methodology cannot incorporate a commitment to reflexive accountability because it requires that theorists are effectively theorizing from nowhere and thus already impartial.

Lastly, an ideal methodology in the case of shared agency merely treats deviations from the model as failures of the non-ideal world to match its ideal. An intersectional approach (and a non-ideal approach more broadly) necessarily requires models for shared agency to be tested, to see if they actually help fulfill the stated purpose, or, if not, what needs to change. In the case of shared agency, this process of testing is ongoing because our social worlds are highly dynamic. In this way, both an intersectional and non-ideal approach to developing models of shared agency will not be the development of a static model, which is then applied in order to manage. Rather, our experiences within communities are already the site through which we reflect on particular relations of power (and their inter-contextual or global interactions) because they are the sites at which we experience, i.e., live, within particular socio-ontological landscapes. Within particular communities (and their histories), we are able to reflect on particular social practices

and norms, experience from birth existing forms of shared agency, their outcomes, and the ways in which we are shaped – for better or worse – by those communal practices. As Collins argues, communities are the sites at which our actions take on meaning and the sites through which those meanings shift through continued action, whether intentionally directed or not.¹⁶ In this way, what an intersectional approach valuably highlights is that our models for shared agency are already localized, *and* that theorizing shared agency is not a static practice. It is a process by which we come to develop partial models, tailored toward specific ends, and reshape those models through our shared practices of testing their efficacy together. It is a deeply empirical process that acknowledges the empirical nature is not testing of partial models against a universal, static picture of reality, but a process of testing our partial models in a dynamic social reality in which co-forming systems of power shape and shift our socio-ontological landscapes (sometimes quite rapidly and sometimes due to our shared agency projects) and that we must, in our various ways, shift with or in response to it.

This is why theorizing about shared agency should start from a commitment to open-ended model pluralism. We will never arrive at a fully complete model, nor is there a limited set of models adequate to cover the myriad ways in which shared agency manifests and shifts. Constrained model pluralism, by contrast, simply focuses on what shared intentions and joint actions “really” are without allowing for any possibility of difference based on context or any possibility of what they could be. A large-scale, anti-oppression protest likely requires a different model than forms of shared agency between a parent and child. This goes further: understanding how forms of shared agency operate with a family structure, for instance, requires attention to background conceptual schemas that inform their understanding of ‘family.’ Partial models that

¹⁶ Collins, *Intersectionality as Critical Social Theory*, 157-188.

capture how shared agency functions on a nuclear family model will look radically different from how a family operates within a kibbutz (especially those in which parenting is communal, not nuclear).¹⁷ That is, partial and dynamic models for shared intentions in a highly individualistic society contain different idealizations than that of a highly communal society. In the event, we require a partial model to start figuring how people from highly communal and highly individualistic societies can successfully share agency for particular purposes, the idealizations in question must be able to account for these differences. In this way, open-ended pluralism presses theorists to further contextualize the purpose of their model as well as continue testing the partial model through action itself. Modelling is thus a continuous process, such that our models are necessarily always partial (i.e., never complete), reflecting what works within a particular case as well as shifting as needed.

V. Closing the Gaps Between Manager and Managed

Since our partial, localized, and dynamic models for shared agency are a second-order form of social reality management, there are still critical questions about the relationship between the managers and the managed. Drawing on particular discussion about the relationship between colonial and decolonial forms of feminism, I show how open-ended model pluralism, one oriented necessarily in terms of social justice, is also a demand to ensure that the gap between the manager and the managed is significantly narrowed or closed in many but not all cases.¹⁸ It is also a demand for a critical interrogation between parochial models. Two

¹⁷ Of course, the purposes of such models would likely be different depending on whether theorists are concerned with reinforcing that family model, calling it into question, or take some other purpose as their end.

¹⁸ There are interpersonal cases, such as a parent-child relationship, in which, for reasons related to child development, certain forms of paternalism are justified. There are likely larger-scale cases that demand more complex analysis and dependency conditions, so I will avoid those here.

incompatible models for community are not necessarily a problem if those models are locally useful. It becomes a problem when: (1) one community or connected set of similar communities tries to universalize their models (especially static and unrevisable ones) in a manner that unjustly exclude other models, (2) when the two communities (and their members) want to undertake a specific action, or (3) if they want to create community together (and to what extent they can or want to preserve features of their partial and localized models in doing so).¹⁹ Structural oppression, again, further compounds this problem. Sometimes the creation of one community is predicated upon the destruction of another. Sometimes the attempt to universalize a model comes from those who largely benefit from forms of oppression, who fail to see that their model is not actually universal (and reflects their particular social-structural positioning) and that the push to universalize is a demand of assimilation toward those who largely do not benefit. Sometimes attempts at solidarity between oppressed communities requires wrestling with the ways in which the drawing of those communal boundaries has been done at the expense²⁰ of those who exist within some or all of those communities;²¹ those who must navigate the exclusions of those

¹⁹ This is to say that two communities operating off of two different models for community might either (1) integrate the two communities and effectively create a new community in doing so, or (2) create a third model for community through which shared community is possible while having reason to preserve the two previous communities and their communal models. Of course, this way of talking about it is already to idealize forms of community as isolated or as being unaffected by the process of (1) or (2). I would also note that separatist models for communities would prevent (1) and (2) but the relationship between separatist and non-separatist communal models that presuppose a shared background, like identity, is already porous. For example, communal models for lesbian separatists would likely prevent the possibility of certain kinds of shared intentions or joint actions with non-lesbian outsiders but interactions between separatist and non-separatist lesbian communities will have a bi-directional effect on both. This is relevant because while lesbian separatists might not share intentions with men in political contexts, the influence of their separatist model (or the forms of social reality management that result from that model) might appear in the ways in which non-separatist lesbian communities engage in shared undertakings with non-lesbian communities.

²⁰ For example, the way community or identity boundaries are drawn as a divide-and-conquer tactic that, in fact, reshapes the socio-ontological landscape.

²¹ This reflects both the ways in which identity communities themselves can draw these lines (for example, the idea women's communities are or should be *purely* about gender) and the ways in which the operations of oppression draw these lines through communities – be it the logic of ‘divide and conquer’ between preexisting communities or the creation of a new taxonomy in which those who previously shared identity are forced to see themselves as having new and separate identities.

boundaries (e.g., you can be X or Y, but not both; or Xs are not Ys, so if you're an X, then you are mistaken about being a Y); and the demand that those who live at multi-axis sites of subordination must choose a side or prioritize one identity community over another.²² Thus, the following discussion shows what happens when a single and supposedly universal model for shared agency is applied in a universal manner. It shows, as I have argued throughout, why we ought to be critical of both idealizations that effectively ignore structural oppression and supposedly universal models; that they tend to uncritically adopt certain conceptual schemas and values, which tend to reinforce forms of structural oppression in their application and become the normative standard through which we manage both the current formations and future possibilities of shared agency.

In *Decolonizing Universalism*, Serene Khader argues that the binary between universalism and relativism frames critiques of universalism as necessarily 'relativist.'²³ Focusing on anti-imperialist, feminist critiques of liberalism and other Western values (like individualized forms of autonomy), the binary between universalism and relativism in moral and political philosophy further inscribes what methodologies and models are seen as legitimate by framing critiques of universalism,²⁴ as Shenila Khoja-Moolji writes, as "regressive, premodern and against the principles of humans rights."²⁵ In this way, those who model features of social

²² For example, in "Learning from the 60s," Audre Lorde expresses the problem in having to "choose between various aspects of [her] identity," given that certain strands of both women and Black liberation movements within the U.S. insisted on the priority of gender over race or race over gender, respectively – and both often excluded lesbians altogether. Lorde cautions against this kind of prioritization that arranges forms of oppression into a "false hierarchy" arguing famously that: "There is no such thing as a single-issue struggle because we do not live single-issue lives." (Audre Lorde, *Sister Outsider*. Berkeley: Crossing Press, 1984/2007: 137).

²³ Serene Khader, *Decolonizing Universalism: A Transnational Feminist Ethic* (Oxford: Oxford University Press, 2019).

²⁴ Khader, *Decolonizing Universalism*, 21-49.

²⁵ Shenila S. Khoja-Moolji, "The Making of Humans and Their Others through Transnational Human Rights Advocacy." *Signs* 42, no. 2 (2017): 377.

reality differently from or in critique of a Western framework are subject to a form of epistemic and investigative gaslighting.²⁶ As Khader writes, “The view that liberalism is the only defensible normative viewpoint makes it difficult to imagine the possibility that one may criticize liberalism, or elements of it, without being a relativist.”²⁷ That is, when a particular set of values (like those found within the tradition of liberalism) serve as the background from which theorists develop models of social reality – and in this case, shared agency – any criticism of those models about the values they presuppose (or merely pointing out that they do, in fact, presuppose those values to begin with) is either taken to be relativistic (and thus reducible to the absurdity of hyper-particularism) or as critiquing a value set that the theorists assumed was, in some general way, above critique. In this way, Khader argues that Enlightenment liberalism presupposes a teleological narrative in which the West positions itself as the furthest along in the path toward justice monism²⁸ and modernity; thereby making itself both the global standard by which to judge moral, political, and social progress in the world *and* justifying its colonialist intervention into any society that is “lagging behind” or rejects some or all of its values.²⁹

²⁶ By ‘epistemic and investigative gaslighting,’ I mean that the terms of this debate, as Khader notes, position those that question universalism (particularly by showing how parochial values and metaphysical assumptions are paraded as universalism) as rejecting or questioning *the only reasonable position available*, such that questioners are discounted and/or treated as being unreasonable (or irrational) from the start.

²⁷ Khader, *Decolonizing Universalism*, 16.

²⁸ Khader argues that justice monism, in the case of feminism, is the claim that “only one type of social or cultural form can house gender justice” and is akin to Amartya Sen’s (2009) critique of John Rawls’ transcendental institutionalism, i.e., “the idea that there is one possible sort of just social arrangements” (30). For Khader, “moralists treat transnational political action as a sort of theatre for sweeping claims about right and wrong rather than a terrain in which practical considerations are at play and power is exercised” (33). What is noteworthy here is that Appiah’s account of ideals and idealization in the context of political philosopher also follows Khader and Sen’s concern. That is, Appiah argues that Rawls’ model for justice misuses idealizations and thus provides an account of justice in a counter-factual world. See: Appiah, *As If*, 115-126.

²⁹ Khader, *Decolonizing Universalism*, 25.

Khader connects this teleological narrative to the specific way in which Enlightenment liberalism frames freedom as a value.³⁰ Enlightenment freedom is a form of positive freedom in which the agent is externally and internally free from the dictates of tradition, particularly religious tradition. Liberal feminism endorses Enlightenment freedom by both (1) positioning the West as more advanced in undoing patriarchy insofar as patriarchy is necessarily bound up in tradition and religious (which is heavily rejected by values associated with Enlightenment freedom, such as the capacity for an individual to reason autonomously), and (2) failing to see Western values (i.e., Enlightenment values) as parochial. On this model, the Enlightenment values are regarded as 'free' from culture and tradition and therefore naturalized as universal truths (or goals toward which all agents and cultures should strive).

Khader argues that Western feminism operates as a form of missionary feminism when it effectively tries to rescue non-Western women from their cultural context and save them into its project of justice monism. That is, Western forms of feminism (largely shot through with whiteness) try to manage the social reality of others under the assumption that they are justified in doing so and/or that this management will, in fact, reshape power relations to the benefit of those whose lives are now being externally managed. Drawing the work of Lila Abu-Lughod and Uma Narayan,³¹ Khader argues that projects which purport to save a person or groups of persons from a particular situation also entail that the person or group is being saved *to* another situation. In the case of missionary feminism,³² Khader notes that Western attempts to 'save' Muslim women from their culture (insofar their cultural backgrounds are framed as necessarily

³⁰ Khader, *Decolonizing Universalism*, 76-98.

³¹ Lila Abu-Lughod, "Do Muslim Women Really Need Saving?" *American Anthropologist* 104, no. 3 (2002): 783-790; Uma Narayan, *Dislocating Cultures: Identities, Traditions, and Third-World Feminism* (New York: Routledge, 1997).

³² Khader, *Decolonizing Universalism*, 21-49.

patriarchal by the West) are simultaneously attempts to save Muslim women *to* the Western world and its values (under the false assumption that the West is more advanced in its progress against patriarchal social orderings).³³ This comes from the assumption that a particular socio-ontological landscape is more just than another and the justification for such a claim comes from appealing to a universal and ideal conception of social and political justice.

For example, one of the major political justifications for U.S. intervention in Iraq during the early 2000s was a concern about women's liberation wherein liberation was equated with Westernization. This is why, as Haifa Zangana argues, that the United States' intervention into Iraq under the pretense of "saving" Iraqi women was devastating to the prewar gains of Iraqi feminism.³⁴ In this case, the attempt by Western and non-governmental organizations (N.G.O) to implement the particular content of Western feminism in Iraq was not only ineffective but damaging to the lives of Iraqi women. Zangana refers to this top-down implementation as colonial feminism.³⁵ The invasion itself undermined decades of local feminist organizing in the country by pretending it did not exist in the first place and inserting so-called feminist N.G.O.s into the region in a top-down attempt to bring in the framework of Enlightenment liberal feminism to help "liberate" Iraqi women. In this way, the assumption is that one partial and localized model for shared agency in cases of resistance is easily exportable to any context without significant revision. One can compare this to the claim that the most effective way to

³³ Uma Narayan, "Sisterhood and "Doing Good": Asymmetries of Western Feminist Location, Access and Orbits of Concern," *Feminist Philosophy Quarterly, Special Issue: In the Unjust Meantime* 5, no. 2 (2019). Narayan notes that there is a fundamental asymmetry in practices of "doing good" or "saving." For example, much of the global justice literature in philosophy focuses on the obligations of Westerners (or those in the Global North) to help non-Westerners, or those in the Global South, but there is basically no discussion of the reverse.

³⁴ Haifa Zangana, "Invading Iraq," in *City of Widows: An Iraqi Woman's Account of War and Resistance* (New York: Seven Stories Press, 2007), 82-107.

³⁵ Zangana, *City of Widows*, 92.

develop models for shared agency comes through the communities themselves and the practices of testing these partial models together through shared action.

What Khader's analysis of transnational feminism shows is how a particular set of values can inform a political model (or a set of models that are often used jointly). In this case, justice monism relies on a single model for justice, one whose values and conceptual schemas are parochial but presented as the only reasonable ones (even when there is allowance for a "reasonable pluralism"). In this way, idealized models for shared agency can function as a form of agency monism. Agency monism captures a limited set of parochial values (i.e., individualized autonomy) and the particular conceptual schemas in which they and the related conception of agency arise, making it easier to define the terms of success or failure prior to investigation how shared agency is understood and expressed within particular contexts. Agency monism also appears in models that do not intend a necessarily universal application but rely on idealizations which presuppose that any contextual differences are largely irrelevant to developing the model (and thus can be mixed back in later when the model is applied).

Lisa Schwartzman's feminist critique of how liberal theorists' use of idealizations like abstraction captures this dynamic well:

Theorizing does not proceed by bracketing all questions about actual society, but rather begins with an examination of power relations in an attempt to understand, criticize, and ultimately change these arrangements. Of course, many liberals would agree that understanding specific issue such as rape or sexual harassment requires studying actual social practices. The difference between my proposal and liberalism is that under my approach, an examination of the structures of power – and the recognition that there *are* structures of power, oppression, and inequality – must be the starting point of the *theory*. In contrast, liberals generally attempt to separate out the application of a theory from its conceptual origins, acknowledging structures of power primarily at the level of application.³⁶

³⁶ Lisa Schwartzman, *Challenging Liberalism: Feminism as Political Critique* (University Park, Pennsylvania: Pennsylvania State Press, 2006), 169.

This is the key difference between how oppression theorists use idealizations in comparison to the idealizations used in the three models of shared agency. Oppression theorists use idealizations that start from non-ideal conditions, idealizing in contextualized ways to examine more localized and globalized functions of power and oppression. The idealizations used by Bratman, Kutz, and Gilbert, however, are high-decontextualized, start from an ideal world, but specifically one understood by a parochial metric of ideal. When such models provide an account of what shared agency is, they do so in accordance with their parochial, over-idealized starting point, which thus excludes, for normative reasons, any forms of shared agency that differ. Moreover, these models function as a bad radar insofar as the predetermined success conditions either label forms of shared agency that do not match the model as failures or the model will fail to detect forms of shared agency that differ substantially (that is, such forms are not labeled as failures, but rather the model fails to register them at all). Instead, success conditions should develop in contextualized ways alongside the development and testing of partial models.

This is why the problem is both about the emphasis on constrained model pluralism *and* the kinds of idealizations used. Even if Bratman, Gilbert, and Kutz models were modified in such a way that they could be incorporated within open-ended model pluralism, their idealizations still appeal to parochial values and conceptual schemas while at the same time erasing the parochiality under a guise of agent-neutrality. What's valuable about Khader's analysis of the failures of missionary feminists is that it highlights why idealizing from particularity results in models that provide a better understanding of the limited scope in which they apply *and* prevents the top-down, unjustified application of models of shared agency onto others in ways that can actually be harmful to forms of shared agency they have themselves developed.³⁷

³⁷ I understand that moral and political theorists will likely raise the question of whether imposing a model for shared agency in this way is ever justified. I am not interested in answering that question, but I will note that my

In conjunction with the *Signs* conference example in Chapter IV, Khader's analysis also highlights why the idealization of agent-neutrality cannot be simply replaced with single-axis idealizations of agential identity. The failures of the conference organizers and the failures of missionary feminisms can only be understood through an intersectional theory of oppression, which tracks, at minimum: (1) a social-ontological claim about the relationship between contextual and dynamic systems of power, social categorization and identity, and communities as the sites through which people experience particular intersections between different systems of power; (2) a methodological and epistemological claim about the value of multi-axis, anti-subordination forms of social theorizing; (3) a set of epistemological practices that places a necessary connection between experience and knowledge, as well as frames social action itself as a source of knowledge.

Chandra Mohanty's work on transnational feminism also captures well how a single-axis, or over-generalized multiple-axis idealizations, can cause similar problems to that of agent-neutrality. Mohanty explains how the identity idealization of 'Third World Women' functions in relation to the abstract category of 'women.'³⁸ Mohanty argues that Western feminist analyses largely appeal to a broadly Western and liberal conceptual schema for agency, which presents those women who live at multiple subordinated intersections of oppression as lacking in agency

main concern about the top-down application is about the ways in which it intersects with the mechanisms of structural oppression and power. It might turn out that through the creation of a multiplicity of models, there are some models for shared agency that are normatively preferable to others within particular contexts. For example, one read of the conference example in Chapter III is that the problem was not simply that the organizers' model for shared agency was parochial and excluded the very people they aimed to include, but also that there may have been a different partial model that would have been a normatively preferable starting point for theorizing through different frameworks for community while also attempting to develop a form of community together (if desirable to all participants). That is, one might argue that the organizers' conception of community was not merely parochial but also normatively inferior for developing and being held accountable to bonds of solidarity.

³⁸ Chandra Talpade Mohanty, "Under Western Eyes: Feminist Scholarship and Colonial Discourses," in *Feminism Without Borders: Decolonizing Theory, Practicing Solidarity* (Durham, NC: Duke University Press, 2003), 17-24.

(i.e. both individually and collectively under the moniker of ‘Third World Women’), as mere victims (e.g., of their culture, their religion, or, alternatively, the process of colonialism),³⁹ and as lacking in power (or as defined fundamentally by their powerlessness).

That is, the idealization of agential neutrality ought not be replaced with single-axis analysis, but theorists ought to also be critical of the ways in which certain multiple-axis identities can be used in over-generalized ways and/or in conjunction with parochial values or conceptions of freedom. In this way, missionary feminism attempts to avoid a single-axis identity by referring broadly to ‘Third World Women,’ but still operates from parochial conceptions of freedom and agency, defining successful agency in this case not only as the absence of oppression but also in relation to a narrow set of parochial values about what agency is and, by extension, what constitutes “deformed” or maladaptive agency. Saba Mahmood similarly argues that this practice also results in largely Western and liberal theorists (including feminist theorists) using a single conceptual framework for investigating various forms of agency that arise within and through different contexts.⁴⁰ For example, Mahmood notes that much of liberal feminist theory reinforces this agent/victim binary through a resistance/submission binary.⁴¹ In Mahmood’s case, Islamic women within the Egyptian da’wa movement are either categorized as submitting to a patriarchal religion or as resisting the patriarchal practices within the religion. However, this binary prescribes the terms of agency prior to investigation and thus

³⁹ Uma Narayan, “Cross-Cultural Connections, Border-Crossings, and ‘Death by Culture’: Thinking about Dowry-Murders in India and Domestic-Violence Murders in the United States,” in *Dislocating Cultures*, 81-117. Narayan argues that western feminism often fails to see potential commonalities between forms of gendered violence between the United States and India, arguing that murdered Indian women were killed by their culture, not by the particular expression of patriarchal violence within the country. That is, there is an over-emphasis of difference by western feminists.

⁴⁰ Saba Mahmood, *Politics of Piety: The Islamic Revival and The Feminist Subject* (Princeton: Princeton University Press, 2011).

⁴¹ Mahmood, *Politics of Piety*, 14-17, 24.

impacts its outcome, conclusion, and/or broader analysis. That is, the liberal feminist model for agency is also used as a universal model for forms of shared agency between women.⁴² In Mahmood's case, it does not actually provide a better understanding of the forms of shared agency that arise with the da'wa movement, but merely reflects how the conceptual framework of liberal feminism interprets the lives of these women who, as Mahmood notes, often do not share that very conceptual framework.⁴³

The takeaway, again, is not that idealizations are useless but rather about how we idealize and the further conclusions we draw from those idealizations. Mohanty argues, for example, that one can use descriptive generalizations (which constitute a form of idealization) while avoiding two forms of, what she calls, *methodological universalism*.⁴⁴ The first kind of universalism is a when a theorist uses a descriptive generalization (e.g., how widespread the practicing of veiling in the Middle East is) to make a universal normative claim about the practices (e.g., the frequency of this practice is causally connected to women's oppression). This results in, for example, the development of resistance strategies that do not track differences in context between practices and forms of veiling,⁴⁵ assumes that veiling itself is patriarchal and

⁴² This is comparable to how Bratman extends his individual account of planning agency into his shared account of planning agency.

⁴³ Mahmood, *Politics of Piety*, 10-22, 153-161.

⁴⁴ Mohanty, *Feminism Without Borders*, 33. Mohanty also argues for a third form that I do not detail here.

⁴⁵ Moreover, this kind of approach cannot incorporate the fact that the same practices of veiling will not be experienced the same way by all women within a shared community, such that the approaches that serve to merely eliminate the practice fail to understand the practice itself is not the source of the problem. (Additionally, following Narayan's point about over-emphasis of difference, there is often a failure to see that Christianity and Judaism also have practices of veiling. For instance, religiously conservative Jewish women cover their hair post-marriage, but do so largely with wigs, having assimilated the practice in a way that hides it from outsiders.) A similar problem occurs with American strains of feminism in which certain clothing items, like a high heel, are taken to be either inherently oppressive (and thus to be eliminated) or inherently liberating (and thus necessarily encouraged). The problem is not a shoe; but rather the problem is how the shoe is used in contextually specific ways to punish women for their clothing choices and/or to maintain segregation in a binary system of gender through the use of gendered visual markers (and thereby punishing men or gender non-conforming people for failing to dress accordingly, particularly in connection to forms of homophobia and anti-queerness).

oppression, and thus prescribes a top-down universal push to eliminate what is seen as a necessary causal factor for gender-based oppression.⁴⁶ Mohanty argues that while a descriptive generalization is not a problem in itself, such generalizations often portray a social practice as a single form rather than as a set of relevantly related social practices whose meanings and histories are contextually different (even if there is a more general way in which they are related). This means that when developing a model for some feature of social reality, theorists ought to take care when using descriptive generalizations to avoid universal, normative attempts to manage that feature of social reality, revising the model as necessary to avoid this. However, it does not mean that descriptive generalizations preclude making normative claims altogether. It simply means that managing parts of social reality needs to be contextually sensitive and attuned to the relationship between the manager and managed. In this case, missionary feminism functions as a way to manage what it considers to be an oppressive practice, effectively reproducing a colonial dynamic between the manager (missionary feminism) and the managed (e.g., Muslim women, particularly those in non-Western countries).

Mohanty's second universalism is the universal application of abstract concepts without attention to differences in the meaning and value of those concepts between contexts and without attention to whether an abstract concept holds in all particular contexts. For example, marriage is an abstract concept but, as an abstract concept, it does not provide any information as to how marriage is understood, valued, and practiced within a particular context (or whether it is at all). Again, the claim that 'marriage is a common practice around the world' is a descriptive generalization comes from theorists noting that there is a relevantly similar set of social practices

⁴⁶ This is related to Khader's concern about missionary feminism and Zangana's concern about colonial feminism insofar as the universal claim often results in top-down resistance strategies that assume there is a universal model for resistance that can be applied anywhere and will produce similar results.

across the world, resulting in an abstract concept like marriage – but this cannot be simply reapplied in theoretical analysis in a top-down manner (especially for normative projects).

What this analysis highlights is that the top-down implementation of a parochial model creates an unjustified gap between the managers (those using the model) and the managed (those who are being externally managed by models that do not reflect their social world). In the case of colonialism feminism, particularly transnational attempts to engage in joint resistance projects by Western feminists do not merely re-enforce the colonial dynamic that they purport to be critically addressing, but instead are themselves a new kind of colonial project, one that attempts to justify itself through a particular ontological framework for gendered oppression and a parochial conceptualization of progress. What Mohanty's methodological analysis further shows is that idealizations do not cease to be bad when those idealizing take social justice as their purpose. That is, even the process of idealizing from concrete conditions requires critical attention to the ways in which such idealizations can easily come to function like those of ideal theory.

Section VI. Friction and Forms of Life

My point in taking the reader through this specific discussion within transnational feminism is three-fold: First, I have argued that Bratman's, Kutz's, and Gilbert's static, idealized models for shared agency all risk reinforcing forms of structural oppression by relying on idealizations that fail to attend to the dynamic relationship between structural oppression, agency, and conceptual schemas. Between the above discussion of transnational feminism and the examples of Chapter IV, my goal is to demonstrate that models developed through the constrained pluralism approach characteristic of an ideal methodology rely on similar sorts of idealizations that appear within particular feminist projects in order to show, more concretely, why this approach to theorizing shared agency will not merely risk uncritically adopting the

unjust outcomes of previous unjustified forms of social reality management, but also, in fact, constitute in and of themselves unjustified forms of second-order social reality management.

Given that avoiding unjustified forms of management requires adopting a methodological commitment to social justice, this highlights further what constitutes this goal is also a site of contestation, not a buzzword to be thrown on under the assumption of shared meaning. The cases discussed show how those who already take themselves to be aiming towards forms of social justice within their own practices, including those related to theorizing, can still quite easily replicate and further intrench unjust systems of power and forms of domination. Given the ease at which this happens with forms of resistance or social action aimed at increasing social injustice, it is unlikely that these idealized models of shared agency borne from a constrained model pluralism approach manage to avoid this problem entirely. In this first case, this is reason to critically assess the idealizations employed by any general model for shared agency and take care in how such idealizations incorporate differences. In turn, this is why there needs to an open-ended pluralistic model approach for theorizing shared agency. Idealizing correctly requires attention to context. However, there is a deep need for critical care to ensure that the use of idealizations does not move to the level of universals (either descriptively or normatively) and that they avoid under-emphasizing or over-emphasizing relevant differences.

Second, models for shared agency constitute a second-order form of social reality management and thus theorizing shared agency necessarily requires critical attention to who is justified in managing what features of social reality and why. The tyranny of a single model (or constrained model pluralism) is that it universally imposes parochial values and conceptual schemas. Open-ended model pluralism does not necessary result in the same problem because

questions of who gets to model what and why are themselves open-ended, and perhaps beget a multitude of contextually useful answers rather than a single answer with universal application.

If our process for developing models of shared agency is about both understanding and management, then an open-ended model pluralism decentralizes the power of this management, requiring further critical engagement with questions such as: (1) who has the relevant epistemic advantage for modelling a particular form of shared agency, for what purpose is it being modeled, how is the model tested, and who takes part in testing it; and (2) what is the ethical nature of relationship between the manager and the managed in this case? For example, if feminist consciousness raising is a shared agency practice, then the development and testing out of partial models for this practice will likely follow from those directly involved in the consciousness-raising process itself, allowing participants to dynamically co-manage a shared process together in virtue of their particular values or needs. More broadly, it seems that the relationship between epistemic priority and management ensures a narrower gap between the manager and the managed *and* better hones the purposes of management.⁴⁷ In this way, an open-ended model pluralism potentially allows for a more democratic forms of social reality management.

Third, and finally, theorizing shared agency by co-developing partial models within and between particular communities is often not about merely determining what shared agency “really is,” but rather about learning what current shared practices work or not, revising our models in light of seeing new possibilities for shared social action. Imagine, for example, if we

⁴⁷ That is, it would be concerning if Men’s Rights activists developed models of feminist critical consciousness raising and those models were used to manage the process because the purpose of the management would likely be akin to a kind of sabotage. Thus, in some cases, the relationship between the manager and the managed can be narrowed down to particular forms of community self-management. On the other hand, if there is a significant break between the manager and the managed, especially those marked by a significant difference in social or political power, then this requires further justification.

relied on a model for shared agency that predated all the technological advances of the 20th and 21st centuries, such that its success conditions excluded (or could not account for) forms of shared agency that arose as a result. That model would not only fail to be useful but it also would have failed to capture the new possibilities for shared agency as they were unfolding (which themselves alter the socio-ontological landscape through which the development of new technologies and their uses play out), preventing future-oriented questions about what we wanted from those possibilities, which possibilities we should encourage, and which should be exercised with caution. This, I take, to be perhaps the most important part about the ethics of social reality management in the case of shared agency. The ability of our models to be future-oriented in this way allows us to see what possibilities these dynamic and socially embedded processes of developing partial models gives rise to and what it potentially suppresses if done carelessly or perniciously, and to be able to be proactive in regards to these possibilities – both in terms of understanding the limitations of our models and in managing what the future might look like. An open-ended model pluralism is thus valuable because it allows individual models to assess what possibilities they open up, both on their own and in dialogue with other models.

If Gail Polhaus is correct about the role of community-building in the development of a standpoint, then we are pressed to consider what this means for navigating a multitude of models.⁴⁸ I take José Medina’s account of epistemic resistance to be helpful here.⁴⁹ Epistemic resistance is: “the use of our epistemic resources and abilities to undermine and change oppressive normative structures and the complacent cognitive-affective functioning that sustains

⁴⁸ Gail Polhaus, “Knowing Communities: An Investigation of Harding’s Standpoint Epistemology,” *Social Epistemology* 16, no. 3 (2002): 291-292.

⁴⁹ José Medina, *The Epistemology of Resistance: Gender and Racial Oppression, Epistemic Injustice, and Resistant Imaginations* (Oxford: Oxford University Press, 2013).

those structures.”⁵⁰ Drawing on Ludwig Wittgenstein, Medina argues *epistemic friction* occurs between different language games, or standpoints, and is productive for interrogating points of exclusion or hegemony in or between different games.⁵¹ Epistemic friction is thus critical for *the interaction of resistant imaginations*:

Imaginations with different moral and political sensibilities can function as *epistemic counterpoints* to each other. The interaction of imaginations can provide a venue for moral and political learning. By comparing and contrasting their imaginative resistances, people can become sensitive to other ways of imagining and inhibiting worlds of possible experiences.⁵²

In this case, the interaction of resistant imaginations is a structural phenomenon produced through certain practices of interaction.

What I find valuable here for the question of how and in what ways to model shared agency is the idea that friction itself is productive, which I take as further proof for the value of open-ended model pluralism and its capacity to develop partial models for more massively shared forms of agency, as needed, by critical attention to the frictions between such models. Patricia Hill Collins similarly argues that intersectionality, as a (proto) critical social theory, functions as a site for dialogue (or friction) between different resistance knowledge projects,⁵³ which, given the intimate connections between methodology and a non-idealized epistemology, is necessary for co-developing partial models between different communities without deeply shared practices in cases where such models are needed. For instance, responding to the complex local and global dimensions of injustice surrounding climate change and environmental destruction will likely require forms of massively shared agency in which those forms develop

⁵⁰ Medina, *The Epistemology of Resistance*, 3.

⁵¹ Medina, *The Epistemology of Resistance*, 15-16, 281-289.

⁵² Medina, *The Epistemology of Resistance*, 256.

⁵³ Collins, *Intersectionality as Critical Social Theory*, 116-120, 144-146.

through a dialectical friction between different resistant knowledge projects, ensuring that such forms of massively shared agency are rooted necessarily, for example, in forms of indigenous, anti-capitalist, decolonial, and anti-racist epistemic resistance. If a model for massively shared agency in the case of climate change avoids this route, then (a) it is unlikely to be massively shared, and (2) those forms of shared agency risk replicating the very anti-indigenous, racist, capitalist, and colonial practices that resulted our current climate problems in the first place.

Medina's account, by proxy, also highlights why a Wittgenstein approach to open-ended model pluralism allows for forms of friction in looking toward new possibilities for forms of shared agency. Take the central question at hand: "What is shared agency?" Bratman's, Kutz's, and Gilbert's answer is something like: If we can determine the necessary and/or sufficient conditions of shared agency, then we can more easily determine what does or does not count as shared agency, offering, in turn, a more universal of shared agency. This is a form of monism. However, Wittgenstein's work on language-games offers a different approach to understanding how questions about the nature of shared agency need not be about a top-down refinement of a concept.

Carl Elliott's analysis of Wittgenstein's notion of a language-game in the case of mental illness helpfully captures how one would apply the term 'shared agency,' in this case.⁵⁴ The seeming proliferation of new disorders, increased diagnosis of disorders previously thought to be rare, and the presence of historically or culturally specific disorders have resulted in both philosophical and psychiatric debates about what constitutes a genuine mental illness. Elliott argues that the standard philosophical approach to a question like, "What counts as mental

⁵⁴ Carl Elliott, "Does Your Patient Have a Beetle in His Box?: Language-Games and the Spread of Psychopathology," in *The Grammar of Politics: Wittgenstein and Political Philosophy*, ed. Cressida J. Heyes (Ithaca, NY: Cornell University Press, 2003), 186-201.

illness?” (or, in this case, “What counts as shared agency?”) is to focus on necessary and/or sufficient conditions, and thus provide a definition that not only provides an easy basis for determining whether a particular case counts, but also provides a unifying definition that finds a commonality throughout all particular instances. As Elliott writes, in this form of investigation: “The philosopher serves as a kind of language czar, deciding for the linguistic community the ways in which words can be legitimately used.”⁵⁵ Following Wittgenstein, Elliott argues that in the case of mental illness:

To look at this extraordinary proliferation of mental disorders and see only an opportunity for rule-writing is like looking at a tropical rain forest and seeing only the opportunity to use a lawn mower. The more interesting question is: How has all this extraordinary complexity come about?⁵⁶

In this way, the question of “what counts as X?” starts from the assumption that X has a clear definition in which all instances of X have a common trait and that the philosopher’s task is to find this trait, thereby determining the appropriate application of X to particular instances.

The value of thinking about the term ‘shared agency’ in the case of language-games is that starts from seeing how ‘shared agency’ (or ‘agency’ in general) operates within particular language-games, which frequently overlap such that shifts in use in one language-game can impact use in another. That is, rather than determining whether the varied uses of agency are correct by way of narrowing the definition, we can instead proceed by looking at how ‘agency’ or ‘shared agency’ is used across a variety of contexts, which are constantly shifting. Open-ended model pluralism thus allows us to see differences between contexts – without also having to immediately take a meta-level normative stance about which are “correct.” As Elliott writes,

⁵⁵ Elliott, “Does Your Patient Have a Beetle in His Box?,” 187.

⁵⁶ Elliott, “Does Your Patient Have a Beetle in His Box?,” 187.

A single word may have many different uses in many different language-games, and these language-games are themselves constantly changing: new language-games are constantly coming into existence, while others become obsolete.⁵⁷

In this way, ‘shared agency’ functions like the word ‘game.’ There is no single feature that unites all particular instances – that is, all potential, partial models – but these models share in ‘family resemblances.’⁵⁸ By proceeding in this way, theorists of shared agency can better see both the similarities and differences between models of shared agency, and come to see, as Wittgenstein writes, the “complicated network of similarities overlapping and criss-crossing similarities in the large and in the small.”⁵⁹ This method does not establish hard boundaries for the appropriate use of the term (or narrow success conditions), but rather the boundaries become vaguer (in the sense that there is a fuzzy boundary) or ambiguous (in the sense of attribution) but do so without giving up on boundaries entirely. That is, the boundary between where green turns into blue might be vague, but we do not lose the sense that some boundary between the two still exists. In this way, starting from a commitment to open-ended model pluralism allows for a kind of radical pluralism; that is, not merely the kind of “reasonable pluralism” captured by a single-model or universalist idealizations (particularly ones associated with ontological and political liberalism).

In this way, we can see how shared agency, like a language, arises in relation to a particular form of life and cannot be understood prior to or detached from a particular form of life.⁶⁰ In an imagined dialogue, Wittgenstein writes:

⁵⁷ Elliott, “Does Your Patient Have a Beetle in His Box?,” 190.

⁵⁸ Ludwig Wittgenstein, *Philosophical Investigations*, trans. G.E.M. Anscombe (Oxford: Basil Blackwell, 1958), §67.

⁵⁹ Wittgenstein, *Philosophical Investigations*, §66.

⁶⁰ Wittgenstein, *Philosophical Investigations*, §19. As Wittgenstein writes, “to imagine a language is to imagine a form of life.”

So you are saying that human agreement decides what is false and what is true?”
– It is what human beings *say* that is false and true; and they agree in the language they use. That is not agreement in opinions but in form of life.⁶¹

To investigate shared agency then is to investigate not merely how the term is used within a particular form of life, but how that form of life itself functions. Thus, to ask how shared agency works requires investigation of the form (or forms) of life in which it operates: How is agency understood? What features are taken to be central? How are the boundaries between individual and shared agency understood or constituted? How is agency evaluated? As Elliott writes,

We use words in certain ways for certain purposes, and those purposes are tied up in cultural forms that are temporally and geographically variable... We cannot look for the beetle in the box. Instead, we must look at the form of life in which the language is embedded.⁶²

By starting from an open-ended model pluralism, theorists treat the question of shared agency as one would treat the question of a game: what are the similarities and differences between different ‘games’? If two people are playing different games, how and to what extent can they transition from those starting points to a shared game? How does trying to develop such shared games alter those involved as well as the forms of life and language-games in which they are embedded?

Moreover, people can inhabit different forms of life and different language-games at different times. It would thus be deceptive to think that individual models operate in isolation from one another, even if they are incompatible with one another, because, particularly in the case of oppression, people are often forced to navigate at the points of incompatibility in practice, even if we know theoretically that we cannot, in fact, utilize incompatible models simultaneously. María Lugones’ account of world-traveling (i.e. navigating different forms of

⁶¹ Wittgenstein, *Philosophical Investigations*, §241

⁶² Elliott, “Does Your Patient Have a Beetle in His Box?,” 192.

life with different language-games)⁶³ in the case of agency and structural oppression captures this well. Forms of life impact the development of intentions and the meanings of the actions in question:

If one can remember the intentions of the person one is in the other world and tries to enact them in the other, one can see that many times one cannot do so because the action does not have any meaning or has a very different sort of meaning than the one it has in the other reality.⁶⁴

In the case of oppression, this takes on an additional layer, because:

... The oppressed know themselves in realities in which they are able to form intentions that are not among the alternatives that are possible in the world in which they are brutalized and oppressed.⁶⁵

In this way, forms of shared agency aimed at resistance structural oppression, for instance, might not always be intelligible or possible within certain worlds, which is further reason for a multiplicity of models. A single, universal model can, at most, capture shared agency within a particular form of life, but the function of universality is to capture all forms of life, which is why such models will simply fail to see certain forms of shared agency altogether. The mechanisms of structural oppression compound the failure of the models. Because forms of life exist in some necessary relation to social-structural positionings,⁶⁶ those inhabiting more relatively privileged positionings will often have little reason to travel to non-dominate worlds, while those who inhabit more oppressed positioning will likely be forced to travel between

⁶³ María Lugones, "Playfulness, "World"-Traveling, and Loving Perception," in *Pilgrimages/Peregrinajes: Theorizing Coalition Against Multiple Oppressions* (Lanham, MA: Rowman & Littlefield Publishers, 2003), 77-100.

⁶⁴ María Lugones, "Structure/Anti-structure and Agency Under Oppression," in *Pilgrimages/Peregrinajes: Theorizing Coalition Against Multiple Oppressions* (Lanham, MA: Rowman & Littlefield Publishers, 2003), 57.

⁶⁵ Lugones, "Structure/Anti-structure and Agency Under Oppression," 59.

⁶⁶ This is not to argue that social-structural positioning determines forms of life, but rather to argue there is a necessary connection between the two.

dominate and non-dominate worlds,⁶⁷ which provides further support for using idealizations, like an epistemic standpoint, rather than idealizations that rely on a non-social epistemology. In this way, traveling between different forms of life – some more similar (or familiar) and others more different – is as a function of ease, based on our fluency within any particular game.⁶⁸ This is why, again, models for shared agency that require mutual knowledge, understanding, or expressions of personal readiness fail to appreciate how participants’ fluency within a form of life (and, in the case of ignorance, one’s failure to see or take seriously other forms of life) impacts where such success conditions are met, and how they are met.

By framing the need for a multiplicity of models in terms of Wittgenstein’s language-games, we can see why Medina’s idea of epistemic friction is so valuable for thinking through the questions of possibility. Perhaps seeing the possibilities for any particular model of shared agency requires to put it in a kind of friction with other models; seeing the compatibility or lack thereof, the exclusions of each model, and the parochial values and conceptual schemas on which the model relies. It is through this friction that certain kinds of forms of imagination are possible. That is, answering the question of what we want shared agency to be requires a form of imagination made possible through the frictions between what we already take shared agency to be. Moreover, those who must navigate many different forms of life are already experiencing this friction – again, this becomes a question of epistemic advantage.

⁶⁷ For example: those who work as the cleaning staff for wealthy families often bear witness to the personal lives of those in the house but are often treated as invisible in that world. Members of the family might have intimate conversations in front of household staff as if they are not present. Those who do this work must navigate the world of their employers but cannot navigate it *like* their employers, and must do this while navigating other worlds in which they might feel different degrees of ease and fluency, within their daily lives.

⁶⁸ For further discussion of ease or lack thereof, see: Lugones, “Playfulness, “World”-Traveling, and Loving Perception.”

And this is the important part for shared agency as a form of resistance: there are many instances wherein people who are not operating off of the same or even similar models for shared agency must figure out how to share agency within one another. Sometimes this might mean that one model is given priority, and thus some participants must incorporate themselves within it. That is, there is a reason that I should learn how to play your game or the reverse, and that might also mean that I must give up some control about how I participate in your game. However, determining this requires appealing to both our games, and justifying why one is preferable to the other in this case. However, I think more frequently there are moments where the two of us realize that neither game is sufficient for our goal or that our particular games are not open to just any player, or there is some other reason for us to figure out how to play a game in which we can both participate. What these moments show is that having to navigate the friction between our models in the understanding that neither is sufficient is the very site through which we can develop forms of shared agency that allow people to play a new, shared game – one that is borne from friction and thus operates without erasing the very sites of difference that gave rise to its possibility.

This is important, again, because there are certainly cases where massive shared agency is required.⁶⁹ This is a central problem for liberal political philosophers like John Rawls.⁷⁰ A ‘comprehensive doctrine’ is just another way to say a ‘form of life.’ The appeal to *reasonable comprehensive doctrines*⁷¹ is the result of asking: how do people from many different forms of

⁶⁹ For example, the case of law: Scott J. Shapiro, “Massively Shared Agency,” in *Rational and Social Agency: Essays on the Philosophy of Michael Bratman*, M. Vargas and G. Yaffe, eds. (New York: Oxford University Press, 2014), Yale Law School, Public Law Research Paper No. 581.

⁷⁰ Here I am referring specifically to: John Rawls, *Political Liberalism*, expanded ed. (New York: Columbia University Press, 2005).

⁷¹ Rawls, *Political Liberalism*, 58-66, 243, 253.

life massively share agency, which binds them normatively to the practice? This approach is central to Bratman's own project, which argues that a shared commitment (a form of shared intention) allows people to decide on a "settled, public common ground that serves as a framework for relevant social thought and action" in the case of shared deliberation.⁷² That is, shared commitments, like Rawls' freestanding political structure,⁷³ determine the terms and limitations of shared deliberation, such that appeals to forms of reasoning or reasons not include in the settled, public common ground are excluded as unreasonable. In this case, 'reasonableness' is effectively the ability to agree to a shared form of agency – a freestanding political structure – and those they cannot agree on the basis of their comprehensive doctrine are unreasonable and thus excluded as participants. This is central to why open-ended model pluralism cannot start from predetermined conceptions of what constitutes a reasonable form of shared agency or a reasonable agent. It prematurely and normatively delimits the scope of theorizing shared agency in ways that exclude forms of life that are unreasonable only insofar as they disagree with the terms of the settled, public common ground (and such disagreement itself is due to the fact that the settled, public common ground is never neutral, often reflecting, again, the interests and values of those with relatively more power and privilege within systems of power).

But there is another path toward finding ways to massively shared agency in cases where it is needed and that participants understand it to be normatively binding. Instead of starting from a model of shared agency that insists that those who disagree with the model, because they do not share its parochial values or conceptual schemas are unreasonable or ideologically

⁷² Bratman, *Shared Agency*, 37.

⁷³ Rawls, "The Idea an Overlapping Consensus," in *Political Liberalism*, 133-172.

backwards, and thus excludes them as participants within and/or as genuine critics of the model (as one to be discarded), we can start from an open-ended pluralism about shared agency – from a multiplicity of models – using the points of friction as productive in developing models for massively shared agency without pre-determining what shared agency is, what successful shared agency looks like, who determines the boundaries of participation, and thus who gets to participate.⁷⁴ This is not about building a settled, public, and common framework for shared agency but rather seeing how to develop partial and contextualized models for problems that seem to necessarily require (at least temporary) forms of solidarity within and between different communities.

⁷⁴ This is a nod to Medina's account of resistance, particularly because Rawls' framework for massively shared agency is a function of overlapping consensus, and I want to be clear that this should not be understood as a claim about consensus – overlapping or otherwise – even though I do not have the space to argue for that here. However, my distinction here tracks Margaret Urban Walker's two models of morality: the *theoretical-juridical model* and the *expressive-collaborative model*. I take in part what I am gesturing at here is that given that modeling shared agency is a normative undertaking, her expressive-collaborative model for morality might be valuable in thinking through ways to understanding this process without hinging the outcome on consensus. See: Margaret Urban Walker, *Moral Understandings: A Feminist Study in Ethics*, 2nd ed. (Oxford: Oxford University Press, 2007), 3-34.

Conclusion

I.

I would like to end by briefly discussing the relationship between social reality management, shared agency, and how people build futures together. As I have argued, the socio-ontological landscapes in which we find ourselves embedded are the outcomes of past and ongoing forms of social reality management. These landscapes are thus not to be naturalized. First-order social reality management is about the process of constructing such landscapes. Second-order reality management, i.e., the use of models for directing and creating forms of shared agency, thus is about how we collectively determine who gets to landscape what and the terms of such allowances. In the first-order, these landscapes affect what forms of shared agency arise and how agents who are similarly or differently positioned within these micro- and macro-landscapes of power develop different forms – sometimes compatible and sometimes not. In the second-order, models for shared agency contribute to the contours of the landscaping itself: how ought people engage in this landscaping together and on whose terms, and who is seen as a potential agent with whom one can undertake this shared process in the first place. That is, who are the managers and who are the managed? Social reality management can thus be, at its worst, two-tier tyranny, and better as a democratic process, one fueled by productive frictions, not a notion of consensus that prematurely delimits the deliberative framework.

Thus, first- and second-order forms of management are fundamentally about shaping the contours of these future landscapes in relation to past and present forms. Our models for shared agency differ as much as our visions for the future, raising a similar set of questions: who is a part of this future planning, who will exist in the future and on what terms, whose visions of the

future are heralded and whose are suppressed or denied? As with shared agency, there is not a single future, but micro- and macro-level futures. As inhabitants of the same planet, we broadly share a future in the case of climate change, but even that broadly shared future will not look the same for all of us nor has its past.¹ Incompatible visions of the future can co-exist, but, perhaps more frequently, the terms of one future stand in some kind of opposition to the terms of another. Shared agency, as a form of planning or otherwise goal-oriented behavior, is about how we manage these futures – whether it be the future of a marriage, a neighborhood, a nation-state, a language, or a lifeworld – and thus about how this management has the power to bring some futures into fruition while denying others the ability to envision their future at all.

To highlight this final point, I return to the example of marriage used throughout this dissertation, particularly marriage as a legally regulated form of shared agency, or, in the case of liberal ideal theory, what John Rawls calls more broadly, “the institution of the family.”² As a legal institution, marriage is one way in which a nation-state provides a model for how two (and sometimes more) people can enter into (or be determined as) a shared form of agency. More broadly, legal forms of marriage are used to determine socio-ontological landscapes of race, gender, nationality and nationhood, sexuality, class, citizenship and immigration status, and differential access to sets of resources (both those provided by the state, such as tax breaks or laws that facilitate inheritance of intergenerational wealth, as well as non-state resources, such as the social currency of being in the “right” kind of relationship, e.g., long-term, monogamous, heterosexual, reproductive). Through practices of naming and taking names, it affects our

¹ One example is that there is no shared timeline for what is referred to as the ‘climate crisis.’ As Kyle Whyte argues, the settler-colonial imaginary of climate change posits the future as dystopian, whereas for many indigenous peoples that dystopia already exists and has for a long time. See: Kyle Whyte, “Indigenous Science (Fiction) for the Anthropocene: Ancestral Dystopias and Fantasies of Climate Change Crises,” *Environment and Planning E: Nature and Space* 1, no.1-2 (2018): 224-242.

² John Rawls, *A Theory of Justice, revised ed.* (Cambridge, MA: Belkemp Press of Harvard University Press, 1971).

genealogical landscapes as well: who is named within one's own familial histories, who is not, and how one is seen in virtue of their names and the histories such names invoke.³ Like the necessary relation between legal and illegal economic markets, legally sanctioned forms of marriage interact with non-legal forms: some people are denied access to legal forms, some people do not need or want their marriage to be state-sanctioned, and others might see their models for marriage as fundamentality incompatible with the legally sanctioned forms. To not commit to this kind of shared agency at all, legal or otherwise, is often another way to confer status to others unequally. Consider the connotative differences between 'spinster' and 'bachelor.'

Legal marriage, as a state-sanctioned model for shared agency, is also used as a way to control the reproduction of citizenry. First, by marking who is allowed to legally reproduce and who is supported in reproduction (e.g., whether pregnancy occurs within or outside of marriage, unequal allocation of resources between those who are pregnant or children, etc.). Second, determining citizenship through marriage and thus triggering investigations allows the state to determine whether such marriages are either a "legitimate marriage of love" or an "illegitimate marriage of convenience" – that is, the legitimacy of the shared agency itself is a condition for a path to citizenship. It is also about whose familial bonds are given support to strengthen over time, whose are ignored or as unseen as lesbian relationships throughout history, and whose are intentionally severed in service of others' hostile visions of the future. In this sense, marriage as a legal institution is about providing a model for shared agency, one that takes up or alters the

³ To highlight a few examples: Political and corporate dynasties in America are often about the name: Kennedy, Ford, or Vanderbilt. Naming in the case of ethnic or religious groups is not only about heritage but also how you are seen by members of your own community as well as those outside of it. For example, a Jewish person with an assimilated name affects how Jewish and non-Jewish people alike see them. That is to say, last names can confer access or the denial thereof but also refer to a particular history in one's own family tree, while overshadowing or completely erasing another.

terms of our socio-ontological landscapes, for the purposes of enacting a particular vision of nationhood into the future, directly affecting who is legally and socially allowed to come into existence, making this vision of the future hostile to those who weren't supposed to exist there in the first place.

Consider the deployment of anti-miscegenation laws, for instance. Such laws are about present and future social reality management, aiming to maintain and expound upon the landscapes created by past forms of management in the case of race and gender. Anti-miscegenation laws offer a legal model for what constitutes “the correct” form of marriage, but specifically as a form of social control over present and future peoples. Such laws seek to maintain these socio-ontological contours of gender and race in the present sense of who is allowed to create familial bonds with whom, but more specifically about policing such boundaries in terms of who is allowed to come into existence and on what terms. Such laws also often operate in conjunction with other processes that affect: which groups ultimately benefit from a particular legal vision of marriage; who has been denied access altogether (in the past or presently); who is allowed, encouraged, or forced to reproduce and on what terms; who is denied reproduction (e.g., through laws that render their children illegal, forced sterilization, or lack of access to adoption); and who is allowed to physically reproduce but denied the ability to raise their children within their own lifeworld (e.g., the separation of children from parents for the purposes of education and socialization, which often weaponizes practices of adoption to break up familial relations, both nationally and internationally). In this way, anti-miscegenation laws highlight broader violent and unjustified forms of social reality management: eugenics, genocide, and colonization.

Genocide, eugenics, and colonization are the clearest forms of unjustified and tyrannical social reality management. Genocide is not only about erasing the physical existence of a group of people, but rather is an attempt at existential erasure: to wipe the memory of their existence from the physical, social, and historical record. Similarly, colonization is about globalizing a parochial vision of who gets to manage the terms of social reality, who is managed, and the narrative of management itself, often through the use of genocide and eugenics. Eugenics takes on a particular vision of betterment, directly attempting to alter the biological and pseudo-biological terms of human existence, under the assumption that certain lives simply are not worth living (e.g., gene editing to prevent certain forms of human existence),⁴ inherently dangerous, or merely not good enough. Like colonization, genocide and eugenics projects are not an event, but a structural process, marked by overt and covert forms of violence across time.⁵ For the survivors in this ongoing process, our current socio-ontological landscapes function as hostile architecture, designed to exclude those regarded as unwanted, as having less value, or those who were not supposed exist in virtue of the design itself. To survive is also to bear witness to denial: the denial that allowed it to happen, the denial of its happening, and the denial of its continued but sometimes unrecognizable happenings.⁶ Those who ‘survived’ in virtue of existing at all in the

⁴ For example, the search for a ‘gay’ gene or a ‘fat’ gene is a eugenics project, often focused on the ability to eliminate the existence of such people by limiting, for instance, diversity of body size. See: Kathleen LeBesco, “Quest for a Cause: The Fat Gene, the Gay Gene, and the New Eugenics,” in *The Fat Studies Reader*, eds. Esther Rothblum and Sondra Solovay (New York: New York University Press, 2009), 65-74.

⁵ Patrick Wolfe, *Settler Colonialism and the Transformation of Anthropology* (London and New York City: Continuum International Publishing Group, 1998), 2. Wolfe discuss this in the case of settler-colonialism: “The colonizers come to stay—invasion is a structure not an event.” However, I think it is more broadly helpful to consider other forms of colonialization as well as genocide and eugenics projects as structural processes rather than limited events in time.

⁶ Michelle Alexander argues that white supremacy, as a structural form of oppression, maintains its power in “preservation through transformation.” That is, oppression is adaptive such that its new formations might be mistaken for social and political progress when, in fact, unjustified forms of power have shifted to a new, harder-to-recognize form, which may or may not constitute some level of progress in relation to past forms, but often results in some people overstating the extent to which moral or social progress has occurred. See: Michelle Alexander, *The*

face of past processes of erasure are still surviving current forms of erasure. In this way, violent forms of social reality management result in different timelines,⁷ particularly in the sense of social, political, and moral progress,⁸ but also the present sense of history. For oppressor groups, such violence is seen as a limited action that occurred long ago, but for the oppressed, histories of violence feel close and memories still vivid, in part because these histories are the material through which the present and future are shaped. Mismatch in perceptions about moral, social, and political progress shows why models of shared agency must attend to the contours of our socio-ontological landscapes. No shared forms of resistance can be born if there is not some agreement about where we are. Shaping futures in a shared way requires mutual understanding about the terms of the present. It's not simply that going 'there' requires knowing where 'here' is, but rather 'here' determines where 'there' is.

Put bluntly, there are many who exist today who were never supposed to exist. There are ways of regulating the terms of shared agency that better ensure such people are denied the ability to manage, or suppressed in their successful attempts to manage, as evidenced by

New Jim Crow: Mass Incarceration in the Age of Colorblindness, revised ed. (New York, NY: The New Press, 2012), 21.

⁷ This is not to ignore that time itself is not shared. While many must live under a Gregorian calendar coupled with a standardized measurement in which time itself is linear, these forms of time do not accord, for example, with lunar calendars in which transitions between days is marked by the vague shifts in sunrise and sunset rather than at the end of a 24-hour time measurement. Thus, the sense of when a new day begins, let alone a new year, is deeply unshared, but forced assimilation to a globally standardized (but still parochial!) can radically alter the time perception of those forced to assimilate. Sense of time can also differ in terms of cultural practices (e.g., if a class lists 3 p.m. as the starting time, when does the class actually start? How are lateness and being on time understood in this case?), personally (i.e., someone who is often late in virtue of their own communally shared understanding of time), metaphysically (e.g., linear, circular, merely illusory or non-existent), and relations between what we refer in analytically separate terms as 'past,' 'present,' or 'future.'

⁸ For example, Chester Pierce details this stark perceptual mismatch in the case of racial progress between Black and white Americans: "In almost any black-white negotiation each participant views things differently, depending on whether he is white (the offender) or black (the offended). For instance, the psychological hallmark of racism is the altogether too well-known tendency for whites to congratulate themselves, before a black, concerning what marvelous "progress" is being made. To the perception of the white offender, this is true and reasonable. From the vantage point of the black offended, however, this is both untrue and unreasonable." See: Chester Pierce, "Offensive Mechanisms," in *The Black Seventies*, ed. Floyd B. Barbour (Boston: Porter Sargeant, 1970), 267.

predictable social and political backlash to anti-oppression resistance. Even when genocide, eugenics projects, and colonization are not ultimately successful in their past attempts at complete erasure, they drastically impact how those targeted imagine and plan for their future, both individually and collectively, including but not limited to, the impact of trauma on agency and imagination, both individually and collectively. To be effective, genocide, eugenics, and colonization all require honed forms of shared agency – models that enhance participants’ abilities to bring about these collective ends. This honing is one that comes from direct attention to how values and ontological assumptions are included within the models, determining what agency is, who counts as an agent and who doesn’t, who can share agency with whom and for what purposes, and what values inform such shared projects. Resistance to these structural forms of oppression is thus about questioning the terms of such shared agency projects and developing forms that allow people to resist these existential threats in pointed ways. Nevertheless, existing together, or flourishing together, can itself be a form of resistance to others’ exclusionary frameworks for sharing agency that function as tyrannical forms of second-order social reality management.

II.

This is ultimately why I have chosen to conclude my dissertation with a discussion about the future, by focusing first on the relationship between violent histories and the present. Shared agency is largely future-oriented: we are trying to bring about some future end together by altering the current socio-ontological landscapes in order to engender it. This is why socio-ontological landscapes function in terms of affordances,⁹ enabling and constraining the

⁹ For discussion of shared affordance, see, in particular: Julian Kiverstein, “Empathy and the Responsiveness to Social Affordances,” *Consciousness and Cognition* 36, no. 1 (2015): 532-542.

possibilities for shared action by shaping the agents who act, the kinds of actions they can undertake together, the meanings of those actions, and the processes of knowledge production that inform, for starters, our differential understanding of these landscapes, particularly how and why we act within them. Our micro- and macro-level visions of the future are thus also shaped in relation to where people, as individuals and as members of social groups, are positioned (and repositioned) with micro- and macro-levels of these landscapes. Theorizing about shared agency, developing partial models that help us to better understand and enhance our capacity to do so, is theorizing about how to build our collective futures, which is why asocial, depoliticized, and static approaches to this necessarily dynamic and power-laden process are ideologically and ethically concerning. Failure to appreciate how an idealized model attempts to separate the model and its creators from the contours of our social world by appealing to a parochial value like neutrality, which is itself a useless and often harmful idealization, is a failure to see how such models themselves contain a picture of the future by offering a normative picture of how we ought to plan for it.

The ability to idealize in this way highlights what kinds of features in our past and current social worlds such theorists take to be worthy of consideration in for planning the future. When the ongoing legacies of violent and unjust forms of social reality management are treated as contingent to these models rather than partially constitutive of them, then these models do not merely risk reiterating forms of ontological oppression. Rather, such models give us a glimpse of how this unjust, future-oriented management likely occurs in the first place, by erasing the histories of these forms of unjust management and denying that the current manifestations of ontological oppression have anything but merely contingent bearing on the terms of the models themselves. This is why oppression theorists have been so rightly concerned with particular

kinds of idealizations. The use of an idealized social ontology and epistemology along with the idealization of agent-neutrality are themselves acts of denial and exclusion. Because our models of shared agency are a form of second-order reality management, the use of these idealizations facilitates further forms of denial and exclusion in our understanding of this process itself, enhancing first-order forms of management that reproduce such practices and the manufactured boundaries they maintain.

III.

It is the intimate connection between shared agency as largely future-oriented and as a first-order form of management and models for shared agency as a second-order form of management that serves as the basis for my ideological and ethical concerns. This is also the reason why I take it that my concerns should not be easily dismissed by the very body of analytic literature that I herein critique, or relegated to the status of mere feminist or non-ideal critique within the literature (especially as one to be *allowed* some accommodation by those positioned more centrally within this subfield and/or treated as a “special interest” problem better dealt with in a different subfield, or in a different field altogether, such as women’s studies or sociology).¹⁰ My critique is not a failure to understand what is or is not within the scope of these individual projects. It is a critique of the scope itself and the methodological basis by which many theorists of shared agency co-determine that scope. I write this because one might respond (as many already have) to my ideological and ethical concerns by arguing that while Bratman’s, Gilbert’s,

¹⁰ In this sense, there might be a clash in interpretative communities within the bounds of philosophy, for starters, but this clash is not reason to dismiss critiques about one’s interpretive community, especially when that critique is derived from philosophical communities whose own academic histories are defined, in part, by having to battle against the exclusionary practices of the discipline itself. For further discussion of this in the case of professional philosophy, see, e.g.: Kristie Dotson, “How Is This Paper Philosophy?,” *Comparative Philosophy* 3, no.1 (2012): 3-29.

or Kutz's models might be problematic or too limited, I have too radically overstated their potential for harm, that these models will not actually be used to manage in the way I have described, or that perhaps I am "seeing oppression where there is none." This is precisely to misunderstand the aim of my critique by focusing on the individual wires of how I have stated the problem, rather than how those wires come together to constitute a cage. As Marilyn Frye argues:

One can study the elements of an oppressive structure with great care and some good will without seeing the structure as a whole, and hence without seeing or being able to understand that one is looking at a cage and that there are people there who are caged, whose motion and mobility are restricted, whose lives are shaped and reduced.¹¹

In this case, failure to address oppressive social-ontological landscapes in developing a model for shared agency is a failure to see how peoples' possibilities for shared agency are impacted by the cage as a whole, not merely the individual wires. Here I suggest another kind of failure of interpretation, namely that my goal in this dissertation is to explain how the use of an ideal methodology and unjustified idealizations in the case of constrained model pluralism forms a kind of cage itself, emphasizing the ideological nature of this broader methodological trend (as shown through particular cases) and why it is ethically concerning. To argue that one particular model might not create conditions for large-scale harm is to focus on the details of one wire and, in some cases, a failure to see the other wires altogether, let alone the cage. Sometimes the possibility of interpretation requires one "to shift the level of one's perception in order to see the whole picture."¹²

¹¹ Marilyn Frye, *The Politics of Reality: Essays in Feminist Theory* (Freedom, California: Crossing Press, 1983), 5.

¹² Frye, *The Politics of Reality*, 5.

Thus, the problem is a broader trend in a body of literature in which many theorists have refused to grapple with the complex and non-ideal world in which they exist,¹³ justifying this refusal by appeal to a methodological process that assumes there is some theoretically neutral starting point, which itself it can be used to dismiss alternative approaches as ‘biased,’ ‘special interest,’ or ‘ideological’ as opposed to ‘impartial.’ Moreover, an ideal methodology does not always stop at the bounds of the literature but shapes how people engage in philosophy as a shared practice. Recent debates about what the futures of philosophy look like, the pointed calls for civility when those historically excluded from or marginalized within the discipline point out its oppressive structures (and the harmful behaviors of those positioned centrally within it), and the demarcation of what philosophy is or should be by those who have never had to justify their position within it all represent a struggle between incompatible visions of the future of the discipline. The idea of a neutral or impartial starting point and the appeal to rationality is a singular and universal instrument function as a way to police who gets to engage in this shared professional practice and on what terms. To label a paper as not philosophy or a person as not a philosopher, in this case, is to impose a particular model for philosophical inquiry as a form of shared agency, labeling other forms of shared philosophical agency as failures or simply as not philosophical at all.

¹³ Besides my focus on predominate models for shared agency in this literature as evidence for a broader trend, I also take as evidence the fact that analytic work on shared agency that does incorporate concerns about idealization and oppression is minimal and relatively recent. See, e.g.: Barbara Fultner, “Collective Agency and Intentionality: A Critical Theory Perspective,” in *The Palgrave Handbook of Critical Theory*, ed. Michael J. Thompson (New York; Palgrave-MacMillan, 2017), 523-546; Eric Chelstrom, “Identities of Oppression: Collective Intentionality’s Seriality Problem,” in *Collectivity: Ontology, Ethics and Social Justice*, eds. Kendy Hess, Violetta Igneski, and Tracy Isaacs (Lanham, MD: Rowman & Littlefield, 2018), 241-264; Christina Friedlaender, “Resisting Oppression Together: Shared Intentions and Unequal Agents,” in *Collectivity: Ontology, Ethics and Social Justice*, eds. Kendy Hess, Violetta Igneski, and Tracy Isaacs (Lanham, MD: Rowman & Littlefield, 2018), 265-289; Tracy Issacs, “What Would a Feminist Theory of Collective Action and Collective Responsibility Look Like?,” in *Collectivity: Ontology, Ethics and Social Justice*, eds. Kendy Hess, Violetta Igneski, and Tracy Isaacs (Lanham, MD: Rowman & Littlefield, 2018), 223-239.

I return now to the point that deliberation itself is a shared process. As Michael Bratman argues, deliberation does not require a shared commitment about its terms, but a shared commitment makes deliberation easier by providing a “settled, public common ground that serves as a framework for relevant social thought and action.”¹⁴ Such shared commitments determine the boundaries of the deliberative process and what reasons to which can or cannot appeal. In actuality, so-called common ground is settled on unequal terms, including whose testimony is taken up as evidence for shared nature of this ground, making the process of deliberation easier for those who have no need to appeal to what is excluded. That’s why this common ground is never neutral; its embedded values and ontological assumptions determine the space of legitimate reasons, not merely the weight we give to those reasons. It frames what is up for debate, what is not, and who is justified in debating what. Social reality management is, in this way, also a process in which we deliberate about our future-oriented intentions, focusing on the broader plan for the future at which such intentional states gesture. Ideal theory and idealized models for shared agency function by prematurely delimiting the scope of these shared deliberative processes in accordance with a parochial common ground. Because such common ground can be taken as given, as natural, as self-evidently correct (as in the case of Enlightenment liberalism),¹⁵ critiquing them from outside of that common ground is a maddening process because such critique is necessarily positioned as always outside of the scope, even as the goal is a critique of the scope. It is positioned as necessarily unreasonable, as bias or ideology, or as unintelligible, even as such critiques point out the flawed nature of those terms.

¹⁴ Michael Bratman, *Shared Agency* (Oxford: Oxford University Press, 2013), 37.

¹⁵ Serene Khader, *Decolonizing Universalism: A Transnational Feminist Ethic* (Oxford: Oxford University Press, 2019), 16.

This is why open-ended model pluralism ultimately requires a rejection of the kinds of methodological practices that result in constrained model pluralism, such as appeals to neutrality and universal claims about the nature of practical rationality and standards for reasonableness, because such practices preclude critical reflection on the values and ontological assumptions reflected not only within these models but also within the constrained approach itself and the terms of its constraints. These types of models for shared agency have the capacity to harm for the reasons I have listed and the methodological practices that result in such models are, in fact, already harmful. This is true both for the idealized models within this literature, but also non-idealized forms that, in virtue of their over-idealizations, still operate similarly. This is why certain practices within feminism about the terms of both what is ‘shared’ and what is ‘resistance’ also constitute unjust forms of social reality management. In highlighting such cases, my goal is to show that merely starting from a non-ideal and power-sensitive approach does not, in itself, fix the problem. This is why an idealized approach that excludes the non-ideal world altogether is, by comparison, even more ethically concerning.

Shared agency is about the kinds of futures we create together, separately, or struggle toward through our clashing visions of future, imagining what the possibilities for shared agency are and what those possibilities afford in shaping present socio-ontological landscapes toward some future. It is about who gets to participate, why, and on whose terms, about how unjustified forms of social reality management ontologically impact our possibilities in drastically unequal ways, and about whose visions of the future will come into fruition through this shared management and whose will not. I’ve used the terms of tyranny and democracy to broadly capture the ways in which co-planning the future is inherently political: it is about who gets to exist, let alone participate, and the terms of our existence as social beings. This is why even

small-scale models for so-called simple actions between a dyad cannot refuse to contend with the complex and unjust realities of our socio-ontological landscapes. This refusal is, in the first case, a form of denial about how one's model already reflects a particular narrative about such landscapes, particularly that structural forms of oppression, of violence, that structure our social worlds, have, at most, a contingent bearing on both the development of such models and terms by which people come to shared agency with others. In the second case, this refusal is a failure to acknowledge that no one exists apart from these socio-ontological landscapes, that there is no neutral starting point, not even theoretically, that the "simple" act of going for a walk together in public is political.

References

- Alexander, Michell. *The New Jim Crow: Mass Incarceration in the Age of Colorblindness*, revised edition. New York, NY: The New Press, 2012.
- Anazaldúa, Gloria. *Borderlands/La Frontera: The New Mestiza*. San Francisco, CA: Spinsters/Aunt Lute Press, 1987.
- Alcoff, Linda Martín. *Visible Identities: Race, Gender, and the Self*. Oxford: Oxford University Press, 2006.
- Alcoff, Linda Martín. "Epistemologies of Ignorance: Three Types." In *Race and Epistemologies of Ignorance*, edited by Shannon Sullivan and Nancy Tuana, 39-57. Albany, NY: SUNY Press, 2007.
- Anderson, Luvell. "Hermeneutical Impasses." *Philosophical Topics* 45, no. 2 (2017): 1-19.
- Appiah, Kwame Anthony. *The Ethics of Identity*. Princeton, NJ: Princeton University Press, 2005.
- . *As If: Idealization and Ideals*. Harvard: Harvard University Press, 2017.
- Ásta. *Categories We Live By: The Construction of Sex, Gender, Race, and Other Social Categories*. Oxford: Oxford University Press, 2018.
- . "Categorical Injustice," *Journal of Social Philosophy* 50, no. 4 (Winter 2019): 392-406.
- Bailey, Moya. "They aren't talking about me..." *Crunk Feminist Collective* (blog), March 14, 2010. Accessed on July 21, 2017: <http://www.crunkfeministcollective.com/2010/03/14/they-arent-talking-about-me/>.
- Barbara Charlesworth Gelpi, interview by Allison Tracy, March 12, 2013, interview part 2 (57-73), transcript, Stanford Historical Society Oral History Program (SCO932), Department of Special Collections and University Archives, Stanford University Libraries and Academic Information Resources, Stanford, CA, <https://purl.stanford.edu/mn590gg7877>
- Bartky, Sandra Lee. *Femininity and Domination: Studies in the Phenomenology of Oppression*. New York: Routledge, 1990.
- Berenstain, Nora. "Epistemic Exploitation." *Ergo* 3, no. 22 (2016): 569-590.
- Bilge, Sirma. "Intersectionality Undone: Saving Intersectionality from Feminist Intersectionality Studies." *Du Bois Review* 10, no. 2 (2013): 405-425.

—. “Whitening Intersectionality: Evanescence of Race in Intersectionality Scholarship,” in *Racism and Sociology: Racism Analysis Yearbook*, edited by Wulf D. Hund and Alana Lentin, 175-205. Berlin: Lit Verlag/Routledge, 2014.

Bratman, Michael E. “Shared Intention.” *Ethics* 104, no. 1(1993): 97-113.

—. *Intentions, Plans, and Practical Reason*. The David Hume Series of Philosophy and Cognitive Sciences Reissues. Palo Alto, CA: Center for the Study of Language and Information, 1999.

—. *Shared Agency: A Planning Theory of Acting Together*. Oxford: Oxford University Press, 2013.

Burman, Åsa. *Power and Social Ontology*. Malmö, Sweden: Bokbox, 2007. (Published under Åsa Andersson.)

Butler, Judith. *The Psychic Life of Power: Theories of Subjection*. Stanford, CA: Stanford University Press, 1997.

Cappelen, Herman. “Conceptual Engineering: The Master Argument.” In *Conceptual Ethics and Conceptual Engineering*, edited by Herman Cappelen, David Plunkett, and Alexis Burgess, 132-151. Oxford: Oxford University Press, 2019.

Cappelen, Herman, and David Plunkett. “A Guided Tour of Conceptual Engineering and Conceptual Ethics.” In *Conceptual Ethics and Conceptual Engineering*, edited by Herman Cappelen, David Plunkett, and Alexis Burgess, 1-26. Oxford: Oxford University Press, 2019.

Carastathis, Anna. “Basements and Intersections.” *Hypatia* 28, no. 4 (Fall 2013): 698-715.

Cartwright, Nancy. *How the Laws of Physics Lie*. Oxford: Oxford University Press, 1983.

—. *The Dappled World*. Cambridge: Cambridge University Press, 1999.

Chelstrom, Eric. “Identities of Oppression: Collective Intentionality’s Seriality Problem.” In *Collectivity: Ontology, Ethics and Social Justice*, edited by Kendy Hess, Violetta Igneski, and Tracy Isaacs, 241-264. Lanham, MD: Rowman & Littlefield International, 2018.

Cheney, Liz. (@Liz_Cheney). “Please @AOC do us all a favor and spend just a few minutes learning some actual history. 6 million Jews were exterminated in the Holocaust...Twitter June 18, 2019, 9:25 a.m. https://twitter.com/Liz_Cheney/status/1140988893627478018

Code, Lorraine. “Is the Sex of the Knower Epistemologically Significant?” *Metaphilosophy* 12, no. 3-4 (July 1981): 267-276.

—. “Taking Subjectivity into Account.” In *Feminist Epistemologies*, edited by Linda Alcoff and Elizabeth Potter, 15-48. New York: Routledge, 1993.

Collins, Patricia Hill. "Learning from the Outsider Within: The Sociological Significance of Black Feminist Thought." *Social Problems* 33, no. 6 (1986) :14-32.

—. *Black Feminist Thought: Knowledge, Consciousness, and the Politics of Empowerment*. New York: Routledge, 1991.

—. "Intersectionality's Definitional Dilemmas," *Annual Review of Sociology* 41, no. 1 (2015): 1-20.

—. *Intersectionality as Critical Social Theory*. Durham, NC: Duke University Press, 2019.

Crenshaw, Kimberlé. "Demarginalizing the Intersection of Race and Sex: A Black Feminist Critique of Antidiscrimination Doctrine, Feminist Theory and Antiracist Politics." *University of Chicago Legal Forum*, 140, no. 1 (1989): 139-167.

Cudd, Ann E. *Analyzing Oppression*. Oxford: Oxford University Press, 2006.

Dembroff, Robin, and Catharine Saint-Croix. "Yep, I'm Gay': Understanding Agential Identity." *Ergo* 6 (2019): 571-599.

Demeulenaere, Pierre. "Where Is the Social?" In *Philosophy of the Social Sciences: Philosophical Theory and Scientific Practice*, edited by Chris Mantzavinos. Cambridge: Cambridge University Press, 2009.

Dennett, Daniel. *The Intentional Stance*. Cambridge, MA: MIT Press, 1987.

Dotson, Kristie. "How Is This Paper Philosophy?" *Comparative Philosophy* 3, no.1 (2012): 3-29.

—. "Conceptualizing Epistemic Oppression." *Social Epistemology* 28: 2 (2014): 115-138.

Douglas, Heather. *Science, Policy, and the Value Free Ideal*. Pittsburgh, PA: Pittsburgh University Press, 2009.

Elliott, Carl. "Does Your Patient Have a Beetle in His Box?: Language-Games and the Spread of Psychopathology." In *The Grammar of Politics: Wittgenstein and Political Philosophy*, edited by Cressida J. Heyes, 186-201. Ithaca, NY: Cornell University Press, 2003.

—. "The Looping Effects of Enhancement Technologies." *The Journal of Bioethical Inquiry*, 16 (2019): 127-131.

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

Friedlaender, Christina. "Resisting Oppression Together: Shared Intentions and Unequal Agents." In *Collectivity: Ontology, Ethics, and Social Justice*, edited by Kendy Hess, Violetta Ignieski, and Tracy Isaacs, 265-289. London: Rowman & Littlefield International, 2018.

—. "Microaggressions and the Problem of Attributional Ambiguity." In *The Routledge Handbook to Social and Political Philosophy of Language*, edited by Justin Khoo and Rachel Sterken. New York: Routledge University Press, *forthcoming*.

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

Fultner, Barbara. "Collective Agency and Intentionality: A Critical Theory Perspective." In *The Palgrave Handbook of Critical Theory*, edited by Michael J. Thompson, 523-546. New York; Palgrave-MacMillan, 2017.

Gessen, Masha. "The Unimaginable Reality of American Concentration Camps." *The New Yorker*, June 21, 2019. <https://www.newyorker.com/news/our-columnists/the-unimaginable-reality-of-american-concentration-camps>

Gibson, J.J. "The Theory of Affordances." In *Perceiving, Action, and Knowing: Towards an Ecological Psychology*, edited by John D. Bransford and Robert Shaw, 127-142. Hoboken, NJ: John Wiley & Sons, 1977.

Gilbert, Margaret. *On Social Facts*. Princeton: Princeton University Press, 1989.

—. *A Theory of Political Obligation*. Oxford: Oxford University Press, 2006.

—. *Joint Commitment: How We Make the Social World*. Oxford: Oxford University Press, 2014.

Hacking, Ian. *The Social Construction of What?* Cambridge, MA: Harvard University Press, 1999.

Haraway, Donna. "Situated Knowledges: The Science Question in Feminism and the Privilege of Partial Perspective." In *The Feminist Standpoint Theory Reader: Intellectual and Political Controversies*, edited by Sandra Harding, 81-101. New York: Routledge, 2004.

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

—. "Introduction: Standpoint Theory as a Site of Political, Philosophic, and Scientific Debate," In *The Feminist Standpoint Theory Reader: Intellectual and Political Controversies*, edited by Sandra Harding, 1-16. New York: Routledge, 2004.

—. *Science and Social Inequality: Feminist and Postcolonial Issues*. Urbana and Chicago: University of Illinois Press, 2006.

Hartsock, Nancy. "The Feminist Standpoint: Developing the Ground for a Specifically Feminist Historical Materialism." In *Discovering Reality: Feminist Perspectives on Epistemology, Metaphysics, Methodology, and Philosophy of Science*, edited by Merrill B. Hintikka and Sandra Harding, 283-310. Dordrecht, Netherlands: Kluwer Academic Publishers, 1983.

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

Held, Virginia. "Can a Random Collection of Individuals be Morally Responsible?" *Journal of Philosophy* 67, no. 14 (1970): 471-481.

hooks, bell. *Feminist Theory: From Margin to Center*. Boston, MA: South End Press, 1984.

Hossenfelder, Sabine. *Lost in Math: How Beauty Leads Physics Astray*. New York: Basic Books, 2018.

Hurley, Susan. *Natural Reasons*. New York: Oxford University Press, 1988.

Isaacs, Tracy. *Moral Responsibility in Collective Contexts*. New York: Oxford University Press, 2011.

—. "What Would a Feminist Theory of Collective Action and Collective Responsibility Look Like?" In *Collectivity: Ontology, Ethics and Social Justice*, edited by Kendy Hess, Violetta Igneski, and Tracy Isaacs, 223-239. Lanham, MD: Rowman & Littlefield International, 2018.

Jaggar, Alison M. "Love and Knowledge: Emotion in Feminist Epistemology." *Inquiry* 32, n. 2 (1989): 151-176.

Jenkins, Katharine. "Ontic Injustice." *Journal of the American Philosophical Association* (forthcoming).

Kant, Immanuel. *Critical of Pure Reason*. Translated by Marcus Weigelt. New York: Penguin Books, 2007.

Keller, Evelyn Fox. *Reflections on Gender and Science*. New Haven: Yale University Press, 1985.

Keller, Evelyn Fox and Helen E. Longino (eds.). *Feminism and Science*. Oxford: Oxford University Press, 1996.

Khader, Serene. *Decolonizing Universalism: A Transnational Feminist Ethic*. Oxford: Oxford University Press, 2019.

Khoja-Moolji, Shenila S. 2017. "The Making of Humans and Their Others through Transnational Human Rights Advocacy." *Signs* 42, no. 2: 377-401.

Kincaid, Harold, John Dupré, and Alison Wylie (eds.). *Value-Free Science? Ideals and Illusions*. Oxford: Oxford University Press, 2007.

King, Deborah. "Multiple Jeopardy, Multiple Consciousness: The Context of a Black Feminist Ideology." *Signs* 14, n. 1 (1988): 42-72.

Kiverstein, Julian. "Empathy and the Responsiveness to Social Affordances." *Consciousness and Cognition* 36, no. 1 (2015): 532-542.

Kuhn, Thomas. *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press, 1962.

Kutz, Christopher. *Complicity: Ethics and Law for a Collective Age*. Cambridge, UK: Cambridge University Press, 2000.

—. "Acting Together." *Philosophy and Phenomenological Research* 61, no.1 (2000): 1-31.

Lawson, Bill E. "Moral Discourse and Slavery." In *Between Slavery and Freedom: Philosophy and American Slavery*, edited by Howard McGary and Bill E. Lawson, 71-89. Bloomington: Indiana University Press, 1992.

Levins, Richard. "The Strategy of Model Building in Population Biology." In *Conceptual Issues in Evolutionary Biology*, edited by Elliott Sober, 73-90. Cambridge, MA: MIT Press, 1984.

Lewis, David. *Convention*. Cambridge, MA: Harvard University Press, 1968.

—. "Psychophysical and Theoretical Identifications." *Australasian Journal of Philosophy* 50 (1973): 33-36.

—. "Logic for Equivocators." *Noûs* 16, no. 3 (1982): 431- 442.

LeBesco, Kathleen. "Quest for a Cause: The Fat Gene, the Gay Gene, and the New Eugenics." In *The Fat Studies Reader*, edited by Esther Rothblum and Sondra Solovay, 65-74. New York: New York University Press, 2009.

List, Christian, and Philip Pettit. *Group Agency: The Possibility, Design, and Status of Corporate Agents*. Oxford: Oxford University Press, 2011.

Longino, Helen. *Science as Social Knowledge*. Princeton: Princeton University Press, 1990.

Lorde, Audre. *Sister Outsider: Essays and Speeches*. Berkeley, CA: Crossing Press, 2007.

Lugones, María. *Pilgrimages/ Peregrinajes: Theorizing Coalition Against Multiple Oppressions*. Lanham, MA: Rowman & Littlefield Publishers, 2003.

MacKinnon, Catharine A. *Toward a Feminist Theory of the State*. Cambridge: Harvard University Press, 1989.

Mahmood, Saba. *Politics of Piety: The Islamic Revival and The Feminist Subject*. Princeton: Princeton University Press, 2011.

McMullin, Ernan. "Galilean Idealization," *Studies in History and Philosophy of Science Part A*, no. 3 (1985): 247-273.

Medina, José. *The Epistemology of Resistance: Gender and Racial Oppression, Epistemic Injustice, and Resistant Imaginations*. Oxford: Oxford University Press, 2013.

Merleau-Ponty, Maurice. *The World of Perception*. Translated by Oliver Davis. Routledge: New York, 2008.

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

—. "'Ideal Theory' as Ideology." *Hypatia* 20, no. 3 (Summer 2005): 165-84.

—. "White Ignorance." In *Race and Epistemologies of Ignorance*, edited by Shannon Sullivan and Nancy Tuana, 13-37. Albany, NY: SUNY Press, 2007.

Mohanty, Chandra Talpade. *Feminism Without Borders: Decolonizing Theory, Practicing Solidarity*. Durham, NC: Duke University Press, 2003.

Narayan, Uma. "The Project of Feminist Epistemology: Perspectives from a Nonwestern Feminist." In *The Feminist Standpoint Theory Reader: Intellectual and Political Controversies*, edited by Sandra Harding, 213-224. New York: Routledge, 2004.

—. *Dislocating Cultures: Identities, Traditions, and Third-World Feminism*. New York: Routledge, 1997.

—. "Sisterhood and 'Doing Good': Asymmetries of Western Feminist Location, Access and Orbits of Concern." *Feminist Philosophy Quarterly, Special Issue: In the Unjust Meantime* 5, no. 2 (2019). <https://doi.org/10.5206/fpq/2019.2.7299>.

Noah, Trevor. *Born a Crime: Stories from a South African Childhood*. New York: John Murray, 2017.

Noble, Safiya Umoja. *Algorithms of Oppression: How Search Engines Reinforce Racism*. New York: New York University Press, 2018.

Ocasio-Cortez, Alexandria (@AOC). "This administration has established concentration camps on the southern border of the United States for immigrants, where they are being brutalized..." Twitter, June 18 2019, 8:03 a.m. <https://twitter.com/AOC/status/1140968240073662466>

O'Connor, Cailin, Liam K. Bright, and Justin P. Bruner. "The Emergence of Intersectional Disadvantage," *Social Epistemology* 33, no. 1 (2019): 23-41.

O'Neill, Cathy. *Weapons of Math Destruction: How Big Data Increase Inequality and Threatens Democracy*. New York: Crown Publishing, 2016.

O'Neill, Onora. "Abstraction, Idealization, and Ideology in Ethics." In *Moral Philosophy and Contemporary Problems*, edited by J.D.G. Evans, 55-69. Cambridge, UK: Cambridge University Press, 1987.

—. "Ethical Reasoning and Ideological Pluralism." *Ethics* 98, no. 4 (July 1988): 705-722.

Pateman, Carol, and Charles Mills. *Contract & Domination*. Malden, MA: Polity Press, 2007.

Perez, Caroline Criado. *Invisible Women: Data Bias in a World Designed for Men*. New York: Abrams Press, 2019.

Pierce, Chester. "Offensive Mechanisms." In *The Black Seventies*, edited by Floyd B. Barbour, 265-282. Boston: Porter Sargeant, 1970.

Plato. *Crito*. In *Five Dialogues: Euthyphro, Apology, Crito, Menno, Phaedo*, edited by John Cooper and translated by G.M.A. Grube, 45-57. Indianapolis, IN: Hackett Publishing Company, 2002.

Polhaus, Gail. "Knowing Communities: An Investigation of Harding's Standpoint Epistemology" *Social Epistemology* 16, no. 3 (2002): 291-292.

Rawls, John. *A Theory of Justice, revised edition*. Cambridge, Mass.: Belknap Press of Harvard University Press, 1999.

—. *Political Liberalism, expanded edition*. New York: Columbia University Press, 2005.

Rich, Adrienne. "Compulsory Heterosexuality and Lesbian Existence," *Signs* 5, no. 4, Women: Sex and Sexuality (Summer 1980): 631-600.

Richardson, Robert C. "Biology and Ideology: The Interpenetration of Science and Values." *Philosophy of Science* 51, no. 3 (1984): 396-420.

Robeyns, Ingrid. "Ideal Theory in Theory and Practice." *Social Theory and Practice* 34, no. 3 (July 2008): 341-362.

Ryle, Gilbert. *The Concept of Mind*. London: Hutchinson and Company, 1949.

Sandoval, Chela. "U.S. Third World Feminism: The Theory and Method of Oppositional Consciousness in the Postmodern World." *Genders* 10 (1991): 1-24.

- Schopenhauer, Arthur. *The World as Will and Representation*. Translated by E.F. Payne. New York: Dover Publications 1969.
- Schwartzman, Lisa H. *Challenging Liberalism: Feminism as Political Critique*. University Park, Pennsylvania: Penn State Press, 2006.
- Searle, John. "Collective Intentions and Actions." In *Intentions in Communication*, edited by Philip R. Cohen, Jerry Morgan, and Martha Pollack, 401-415. Cambridge: MIT Press, 1990.
- . *Mind, Language, and Society: Philosophy in the Real World*. New York: Basic Books, 1998.
- Sen, Amartya. *The Idea of Justice*. Cambridge, MA: Harvard University Press, 2009.
- Shapiro, Scott J. "Massively Shared Agency." *Rational and Social Agency: Essays on the Philosophy of Michael Bratman*, M. Vargas and G. Yaffe, eds. (New York: Oxford University Press, 2014); Yale Law School, Public Law Research Paper No. 581, 2014. Available at SSRN: <https://ssrn.com/abstract=2839482>
- Smith, Dorothy. "Women's Perspective as a Radical Critique of Sociology." *Sociological Inquiry* 44 (1974): 7-13.
- Sperber, Dan, and Dierdre Wilson. *Relevance*. Cambridge, MA: Harvard Press, 1986.
- Strawson, P.F. *Freedom and Resentment and Other Essays*. New York: Routledge, 2008.
- Stemplowska, Zofia. "What's Ideal about Ideal Theory?" *Social Theory and Practice* 34, no. 3 (2008): 319-40.
- Sullivan, Shannon, and Nancy Tuana (eds.). *Race and Epistemologies of Ignorance*. Albany, NY: SUNY Press, 2007.
- Tessman, Lisa (ed.). *Feminist Ethics and Social and Political Philosophy: Theorizing the Non-Ideal*. New York: Springer, 2009.
- . *Moral Failure: On the Impossible Demands of Morality*. Oxford: Oxford University Press, 2014.
- Tollefsen, Deborah. *Groups as Agents*. Cambridge, UK: Polity Press, 2015.
- Tuomela, Raimo. *Social Ontology: Collective Intentionality and Group Agents*. Oxford, Oxford University Press, 2013.
- Vaihinger, Hans. *The Philosophy of "As If": A System of the Theoretical, Practical, and Religious Fictions of Mankind*. Translated by C.K. Ogden. New York: Harcourt, Brace, and Co., 1924.

Valentini, Laura. "On the Apparent Paradox of Ideal Theory." *The Journal of Political Philosophy* 177, no. 3 (2009): 332-355.

Walker, Margaret Urban. *Moral Understandings: A Feminist Study in Ethics*, 2nd edition. Oxford: Oxford University Press, 2007.

Welch, Shay. *Existential Eroticism: A Feminist Approach to Women's Oppression Perpetuating Choices*. Lanham, Maryland. Lexington Books, 2015.

Whyte, Kyle. "Indigenous Science (Fiction) for the Anthropocene: Ancestral Dystopias and Fantasies of Climate Change Crises." *Environment and Planning E: Nature and Space* 1, no.1-2 (2018): 224-242.

Wittgenstein, Ludwig. *Philosophical Investigations*. Translated by G.E.M. Anscombe. Oxford: Basil Blackwell, 1958.

Wolfe, Patrick. *Settler Colonialism and the Transformation of Anthropology*. London and New York City: Continuum International Publishing Group, 1998.

Wundt, Wilhelm. *Ethics, An Investigations of the Facts of Laws of Moral Life*. London: Sonnenschien & Company, 1897.

Wylie, Alison. "The Feminism Question in Science: What Does It Mean to 'Do Social Science as a Feminist'?" In *The Handbook of Feminist Research*, edited by Sharlene Hesse-Biber, 567-578. New York: Sage, 2007.

Young, Iris Marion. *Justice and the Politics of Difference*. Princeton, NJ: Princeton University Press, 1990.

—. *Intersecting Voices: Dilemmas of Gender, Political Philosophy, and Policy*. Princeton: Princeton University Press, 1997.

—. *Responsibility for Justice*. Oxford: Oxford University Press, 2011.

Zangana, Haifa. *City of Widows: An Iraqi Woman's Account of War and Resistance*. New York: Seven Stories Press, 2007.

Zinn, Maxine Baca, Lynn Weber Cannon, Elizabeth Higginbotham, and Bonnie Thornton Dill. "The Costs of Exclusionary Practices in Women's Studies." *Signs* 11, n. 21 (1986): 290-303.

"Editorial." *Signs* 10, n. 4 (1985): 633-636.

"We Are Not Fucking Around." Never Again Action. Accessed on February 24, 2020. <https://www.neveragainaction.com>.