

# Belief Change & Social Change\*

Carolina Flores

## Abstract

Individual beliefs often push back against structural reform, posing an obstacle to social change. Troublingly, beliefs that play this role are hard to change, often resisting counter-evidence. This poses a problem for structuralism, which prescribes structural change without considering how to get individuals to abandon resistant social beliefs. I argue that the structuralist has resources to address the problem of resistant social beliefs. Specifically, I argue that social network change can lead to the abandonment of resistant social beliefs, addressing even forms of active psychological resistance to belief change such as identity-protective reasoning. This solution to the problem of resistant social beliefs has significant implications for the debate between structuralists and individualists. In particular, it shows that careful attention to human psychology and proposing structural interventions are compatible. This makes room for bringing together insights from both individualist and structuralist traditions, allowing for a unified account of the relationship between belief change and social change.

And there's no point sidling up crabwise with a mea culpa look, insisting it's a matter of the salvation of the soul. Genuine disalienation will have been achieved only when things, in the most materialist sense, have resumed their rightful place.

---

Franz Fanon, *Black Skin, White Masks*, xv

## 1 Introduction

People with a criminal record find it hard to get jobs. And, when unemployed, people with a criminal record are more likely to re-offend and end up back in prison. Because a disproportionate number of incarcerated people in the US are Black and Latinx, this vicious cycle contributes to racial injustice. To address this, activists have advocated

---

\***Acknowledgments:** Thanks to César Cabezas, Elisabeth Camp, Nina DeMeo, Jon Ellis, Verónica Gómez Sánchez, Sally Haslanger, Sahar Heydari Fard, Alex Guerrero, Eric Mandelbaum, Annette Martin, James McIntyre, Hanna Pickard, Susanna Schellenberg, Ernest Sosa, Elise Woodard, and Caroline von Klemperer for discussion and feedback on previous versions of these ideas. Thanks also to audiences at LOGOS-Barcelona, Agnes Scott College, and the University of Michigan Epistemology Research Group.

for ‘Ban the Box’ measures, which make it illegal for employers to ask about criminal record in job application forms (Avery and Lu 2021).<sup>1</sup>

This is a structural intervention. Through a change in the law, ‘Ban the Box’ measures change the context in which particular individuals (in this case, employers) make decisions. Structural interventions contrast with individualist interventions, which aim to change individual attitudes as a way to change society. For example, trying to persuade employers to be less suspicious of people with criminal records is an individualist intervention.<sup>2</sup>

Intuitively, the structural intervention seems much more promising than individualist interventions in this case. Persuading employers one-by-one to be less suspicious of formerly incarcerated people is a tall order. In contrast, one would think that employers simply can’t discriminate along these lines if they lack information on criminal records. Unfortunately, ‘Ban the Box’ measures do not seem to help—at least, insofar as the point of such measures is addressing racial injustice. Low-skilled Black and Latino men are marginally less likely to be employed after ‘Ban the Box’ measures than before (Doleac and Hansen 2016). When they are not allowed to ask about criminal record, employers’ beliefs that Black and Latinx applicants generally have a sketchy background kick in and lead them to avoid hiring them. For this reason, in Michelle Alexander’s words, “Banning the box is not enough. We must also get rid of the mind-set that puts black men ‘in the box’ ” (Alexander 2010, 153).

As this case illustrates, individual attitudes—in particular, individual beliefs—can push back against structural change. In the ‘Ban the Box’ case, individual employers’ beliefs lead them to find workarounds to keep up the (racist) *status quo*. In other cases, structural measures lead to individual complacency and moral licensing, with individuals assuming that a few measures mean that fairness has been achieved (Dover et al. 2014, Kaiser et al. 2013). Even worse, we sometimes encounter aggressive individual backlash to structural measures, as has sometimes happened with affirmative action (Hughey 2014).

Given that individual beliefs push back against structural change, achieving social change requires changing beliefs (Madva 2016).<sup>3</sup> This is a notoriously difficult task. As Charles Mills vividly put it about white ignorance, many such beliefs are ones that “resist, fight back...[are] militant, aggressive, not to be intimidated, active, dynamic, refuse to go quietly” (Mills 2007, 13). As a result, even if agents receive counter-evidence to such beliefs, they are not likely to abandon them.

Achieving social change, then, requires us to contend with the problem of *resistant social beliefs*: beliefs that (a) pose obstacles to the success of structural reforms, and (b) actively resist counter-evidence. Such beliefs generate practical difficulties in achieving social change. In particular, they pose obstacles for structuralist proposals: it is not easy to see how a focus on structural reform can succeed in the face of

---

1. Madva (2020) discusses this example.

2. See Ayala-López and Beeghly 2020, Brownstein et al. 2021, and Madva 2020 for more on how to draw this distinction.

3. Two clarifications. First, though the paradigmatic cases of social change in this literature have to do with addressing structural injustice, any kind of change in our social organization is covered by the debate. Second, beliefs are not the only aspect of our mental lives that can push back against structural change. We will need additional interventions to address other aspects.

resistant social beliefs.

In this paper, I will argue that the structuralist has resources to address the problem of resistant social beliefs. Specifically, I will argue that social network change is a powerful structuralist resource for getting agents to abandon resistant social beliefs. This provides a novel defense of the power of structural interventions: they can address even resistant social beliefs.

Importantly, this defense of structural interventions is *methodologically* individualist. It will involve detailed consideration of psychological mechanisms involved in belief maintenance and revision. In this way, this paper pushes back against the tendency to view individualism and structuralism as all-encompassing frames, where one must pick a side and stick to it both in methodology and practical recommendations.<sup>4</sup>

Recommending structural interventions is usually accompanied by hostility towards attending to the nuts and bolts of human psychology (Anderson 2010, Frye 1983, Dixon et al. 2012, Haslanger 2015, Haslanger forthcomingb, Táiwò 2017). Conversely, addressing psychological phenomena is typically taken to require individualist interventions, i.e., interventions that directly target individual attitudes, instead of the context in which those attitudes are produced and maintained (Garcia 1996, Stanley 2015). My argument suggests that such polarization between individualism and structuralism is misguided, opening new avenues for exploring and achieving social change. Practical structuralists should abandon their hostility to methodological individualism, and methodological individualists should broaden their sights to consider structural interventions.

I will proceed as follows. In §2, I will argue that the psychological mechanism of identity-protective reasoning plays a crucial role in sustaining resistant social beliefs. Addressing identity-protective reasoning has often been taken to fall squarely in the province of practical individualism. Against this, in §3, I will argue that social-network-shaping interventions can effectively address identity-protective reasoning. In §4, I will argue that such structural interventions are also powerful when it comes to addressing other factors that sustain resistant social beliefs. Altogether, this supports the view that practical structuralism can address the problem of resistant social beliefs. The upshot is that methodological individualism is compatible with, and can even support, practical structuralism. This makes room for a novel position in the structuralism-individualism debate, combining careful attention to psychology with promoting structural interventions (§5).

## 2 Identity-Protective Reasoning as a Central Mechanism Behind Resistant Social Beliefs

The ultimate goal of this paper is to argue that the practical structuralist has resources to address the problem of resistant social beliefs. This should be of interest to anyone interested in growing their toolkit for social change. But its significance is deeper

---

4. See Haslanger 2020 for the distinction between methodological individualism and structuralism. See also Ayala-López and Beeghly 2020, Brownstein et al. 2021, Davidson and Kelly 2020, and Madva 2016 for discussions of the distinction. Note that I use the term ‘practical individualism/structuralism’ where these other theorists use the term ‘individualism/structuralism about interventions.’

for the structuralist. It is a response to worries that have been taken to check-mate structuralism. In responding to such worries, we can begin to properly understand the power of structural change—not only at institutional and material levels, but also at psychological levels.

Let me explain the concerns for structuralism I have in mind. It is easy enough to grant that structuralists have resources to change many aspects of social structures, such as laws, material conditions, or institutional design. After all, such aspects do not appear to reduce to individual psychology. Laws, material conditions, and institutions persist even when most people think they are unjust, even the ones responsible for enforcing them. In contrast, however, it is not clear how structuralism can change *psychological* aspects of social structures, such as beliefs.

To the extent that structuralists consider how to effect belief change, they have not had much to say about resistant social beliefs. They talk about the distribution of information, what testimony individuals are receiving, and who they trust (e.g., O'Connor and Weatherall 2019, Nguyen 2020, Stanley 2015). But they do not generally consider beliefs that actively resist counter-evidence, or the psychological mechanisms that support such beliefs.

This is an odd gap. Theorists have paid attention to resistant social beliefs in discussing white ignorance, ideological beliefs, and prejudice, among other topics. But they have primarily focused on characterizing these target phenomena, without attending to how they fit into a structuralist picture of social change. This is in part out of an opposition to methodological individualism, i.e. the view that understanding social structures centrally involves looking at the individuals in those structures. This opposition to psychology makes structuralists vulnerable to the objection that they assume a false view of human psychology. Indeed, as Madva (2016) points out structuralists tend to assume “the mirroring view of belief”, according to which beliefs transparently reflect the evidence received. On the mirroring view of belief, changes in social structures inevitably result in corresponding changes in belief, much as changing what is in front of a mirror changes what the mirror reflects.

What does the structuralist have to say about reducing resistance to evidence? Answering this question is the project of this paper. To address this question, I will focus on one source of resistance to evidence that takes prime of place in the social and political domain: identity-protective reasoning. Identity-protective happens when agents attend, interpret, and respond to evidence in ways that enable them to protect cherished social identities (Kahan 2012, Kahan 2015, Kahan 2017). I will then draw on my discussion of this case to consider broader lessons for addressing other sources of resistant social beliefs.

To start, we need to get clearer on identity-protective reasoning, its mechanisms and scope. That is the task of this section. Once this is in clear sight, we will be in a position to consider what causal levers the structuralist might have to address identity-protective reasoning, and thereby resistant social beliefs.

## 2.1 From Social Identity to Identity-Protective Reasoning

Identity-protective reasoning is all about managing one’s beliefs in ways that protect one’s social identities. Social identities (including race, gender, class, religion, political

alignment, sports fandom, and so on) are part of our self-concept (i.e. individuals' sense of who they are; Brown 2000, Tajfel et al. 1979, Tajfel 1982). The social identities that figure in our self-concept are thick with descriptive and normative content. These include both claims about characteristic features of members of the corresponding kind or group, and claims about the norms that apply to members of the group (Knobe et al. 2013), sometimes including norms on what beliefs one should have. To a large extent, we pick up these claims from our culture, but different individuals will have different conceptions of the same social identity.<sup>5</sup>

As self-affirmation theory (Gilbert 2009, Mandelbaum 2019, Sherman and Cohen 2006, Steele 1988) documents, we strive to defend our self-concept. In other words, we want to have a stable and good self-concept.<sup>6</sup> More specifically, we desire to hold on to the view that the features we incorporate in our self-concept are valuable, and to preserve the same features as part of our self-concept.<sup>7</sup> Because the self-concept includes social identities, we strive to defend our social identities. Individuals desire to hold to the claim that e.g. being a man, an immigrant, or a Democrat (with all that the more specific features that this involves for them) is good and valuable, and that they truly belong to these categories.<sup>8</sup>

The desire to regard our social identities as stable and good, in turn, affects how we manage beliefs, and how we respond to evidence that threatens these beliefs. Specifically, it affects how we engage with evidence that threatens beliefs connected with those identities. In other words, we engage in motivated reasoning when it comes to such beliefs: desires to protect the self-concept causally influence how we interact with evidence (Kunda 1990), affecting what we infer from the evidence we have, what evidence is salient to us, and what evidence we gather. By affecting how we interact with evidence, our desires affect what we end up believing. If we lacked these desires, we would interact with evidence differently, and in many cases we would end up with different beliefs.

Summing up, we cherish certain social identities—for example, partisan, racial, national, or professional identities. They become part of our self-concept, of our sense of who we are and of our value. In normal circumstances, we aim to defend both the goodness and stability of our self-concept. Part of protecting our self-concept is defending cherished social identities. And this, in turn, leads us to engage in motivated reasoning when it comes to beliefs that are tied to our social identities. Specifically, we put in effort to find ways to maintain these beliefs in the face of counter-evidence.

---

5. For more on what centering social identities in our thinking and action involves, see [REDACTED].

6. The desires at play in this context are typically not conscious.

7. There are some exceptions to this. People with severe depression (among other mental health conditions) typically have a *negative* self-concept (Tarlow and Haaga 1996). In this case, they don't attempt to defend the goodness of their self-concept (because it is not a positive one), but they do seek to defend its stability (Swann Jr 1992).

8. Other psychological factors conspire to make us strive to defend cherished social identities. For example, desires that those who are not 'on our team' be less competent and deserving—as a byproduct of the desire that 'our side' deservingly wins—can play a role (Klein and Kunda 1992). And there is some experimental evidence that our need to cope with mortality and meaninglessness leads us to reaffirm the distinctive value of our communities (Pyszczynski et al. 2015). See Quilty-Dunn (2020) for discussion of the multiplicity of sources of defensive reasoning.

## 2.2 The Long Reach of Identity-Protective Reasoning

The psychological processes outlined above most transparently show up when beliefs in the goodness of cherished social identities are under attack. For instance, they are liable to be at play when white people are accused of benefiting from racist social structures, an accusation which compromises connections between whiteness and innocence (Sullivan 2006).

But identity-protective reasoning extends much further. We have reason to think it causally sustains many resistant social beliefs that are not directly connected with social identities. For one, a key mechanism by which we support our sense of the goodness of our social identities is by making positive comparisons with salient outgroups (Brown 2000, Tajfel et al. 1979). For this reason, identity-protective reasoning can turn into identity-*attacking* reasoning: reasoning that contributes to holding negative beliefs about salient outgroups.

Further, identity-protective reasoning extends to beliefs that at first sight appear distant from social identities. Here is an example.<sup>9</sup> Meat-eaters often endorse “the four Ns” about meat-eating: they believe that eating meat is (1) necessary (e.g., for protein), (2) natural (i.e., humans are meant to do it), (3) normal (i.e., humans generally do it), and (4) nice (i.e., meat tastes good) (Piazza et al. 2015). It is hard to get meat-eaters to abandon these beliefs. Interestingly, men are less likely than women to abandon these beliefs in the face of evidence, and more likely to respond to arguments against them with defensive reasoning. A promising explanation for this finding is that this defensiveness is a manifestation of identity-protective reasoning. Given culturally dominant connections between meat-eating and masculinity (“Real men eat meat”), men are more likely than other people to incorporate meat-eating into their identity (Rothgerber 2013). As a result, arguments against the four Ns are more likely to be met with defensive reasoning. The desire to defend masculinity under a dominant cultural conception makes (many) men closed-minded about meat-eating.

A more general example of the long-reach of identity-protective reasoning comes from thinking about political partisanship. Defending partisan identities (Van Bavel and Pereira 2018) has the power to affect many of our beliefs. This is because partisan identities are typically associated with a wide range of empirical beliefs. For example, under some current conceptions, being a Republican is strongly associated with believing that climate change isn’t real, that vaccines are unsafe, and that creationism is true (Rutjens et al. 2018). Indeed, an increasingly wide range of views, including *prima facie* non-political views (e.g. about coffee consumption or fashion), have come to be associated with partisanship (Dellaposta 2020). This means that attempting to protect one’s partisan identity often results in defensiveness about very large swathes of a person’s web of belief.

Finally, identity-protective reasoning is long-reaching in that it affects attitudes that are not beliefs. This discussion does not cover only on-off beliefs: it also includes credences and suspension. For simplicity, I will use the term ‘belief change’ to cover abandoning one’s belief that  $p$ , ceasing to suspend on whether  $p$ , changing one’s degree of belief in  $p$ , or coming to believe that  $p$ . Including suspension makes this discussion straightforwardly relevant to addressing active ignorance (Alcoff 2007,

9. See Quilty-Dunn (2020) for discussion of this example.

Medina 2013, Mills 2007). Further, if you think, as I do, that implicit biases are beliefs, my discussion will be relevant for addressing implicit bias.<sup>10</sup>

In sum, as individualists have emphasized, identity-protective reasoning makes a significant contribution to maintaining doxastic attitudes in the face of counter-evidence, including (and perhaps especially) ones which pose an obstacle to social change. Specifically, we have good empirical reason to think that identity-protective reasoning is at play for many of our socially relevant doxastic attitudes. Identity-protective reasoning leads us to maintain such attitudes in the face of counter-evidence. For this reason, we need strategies for addressing identity-protective reasoning to solve the problem of resistant social beliefs.<sup>11</sup>

### 3 Changing Social Networks as a Means to Changing Social Identities

At the beginning of §2, I pointed out that individualists often appeal to identity-protective reasoning to argue that structural interventions are insufficient. They argue that identity-protective reasoning provides a decisive obstacle to practical structuralism. Against this, I will now argue that there are valuable structuralist resources for addressing identity-protective reasoning. Specifically, I will argue that changes in social networks are a powerful means to change the structure of the self-concept.

This matters to the problem of resistant social beliefs. The threat to an agent's self-concept has to be significant enough for them to engage in identity-protective reasoning. Specifically, it must be the case that we sufficiently value that identity, that the belief targeted is sufficiently important to that identity, and that the counter-evidence is strong enough (Lieberman and Chaiken 1992). I will argue that, by changing social networks, we can reshape the self-concept to reduce and redirect identity-protective reasoning.<sup>12</sup>

#### 3.1 Social Environment Systematically Shapes Social Identities

Our social environment constrains and shapes our social identities. It systematically constrains which identities are available, what their content is, and which ones are socially appropriate. This is intuitive. In some environments, women are valued for centering motherhood in their self-conception, whereas in others doing so might be looked down upon. Similarly, people in queer communities have available fine-grained gender and sexual orientation identities, which people in heteronormative contexts do not have available to incorporate in their self-concept.

- 
10. See Mandelbaum (2016) for a compelling case for the claim that implicit biases are beliefs.
  11. Note that this does not mean that identity-protective reasoning *always* poses obstacles to social change, only that it sometimes does. Further, the epistemic and practical rationality of identity-protective reasoning in different contexts are topics that deserve careful analysis (see Kahan 2015 for one proposal), and which interact in interesting ways with debates about externalism and internalism in epistemology (Srinivasan 2020). For the purposes of this paper, all we need is the claim that some instances of identity-protective reasoning contribute to the maintenance of resistant social beliefs.
  12. I will then (in §4) argue that such changes in social networks also reduce the power of other defensive mechanisms that frequently sustain resistant social beliefs.

Self-categorization theory (Turner and Oakes 1986, Turner et al. 1987, Turner 2010, Turner and Reynolds 2011) provides useful resources for understanding and predicting the structure of the self-concept. Whether social aspects of the self-concept, as opposed to individual ones (personal traits, values, and preferences), take a central place in one's behavior is a function of social context. Contexts where there are rigid social boundaries, status differences, and conflict between groups drive individuals to center social (as opposed to individual) identities in their self-concept (Turner et al. 1987). This suggests that reducing social boundaries and creating more egalitarian forms of social organization can reduce how much agents center social identities.

Specifically, we categorize collections of individuals as forming a group to the degree (*inter alia*) that the perceived differences between them are less than the perceived differences between them and other people (outgroups) in the relevant context of comparison (Turner and Oakes 1986). What we consider our ingroup, and, correspondingly, which identities we take up, is deeply dependent on who happens to be around us, and the relationships between them.<sup>13</sup> Finally, social context shapes the *content* of the social categories at play. Who counts as a prototypical member of a group, and what traits are taken to be characteristic, depends on the contrast class at play (Hogg et al. 1990), i.e., on who is the outgroup in the context.

None of this is to say that which social identities we center is *determined* by the people around us. Different people can and do behave differently in the same environment, drifting towards different social groups and adopting different social identities. There is an interplay between individual agency and the environment, leaving individuals space to embrace or reject social identities that are put onto them.<sup>14</sup> For the purposes of defending the role of structural interventions in shaping the structure of agents' self-concepts, what matters is that there are some (defeasible and context-sensitive) generalizations about how individual social identities are causally affected by social context—which there seem to be, if self-categorization theory is along the right lines.

Offering further support to the idea that social identities at the individual level are shaped by social context, the *common ingroup identity model* (Gaertner et al. 2000) indicates that cross-group contact (contact among distinct social groups in conditions that foster respect and a sense of community) often produces new shared group identities among participants. These identities can fully replace pre-existing ones. For example, in successful company mergers, people stop identifying with the smaller company for which they used to work, and adopt a new identity associated with the larger company (Giessner et al. 2012). New identities can also sit alongside pre-existing identities: think here of how college students who feel a sense of school pride often adopt 'student at University X' as part of their sense of self, alongside existing religious, ethnic, or political affiliations (Dovidio et al. 1998).

The key point in the common ingroup identity model is that *cross-group contact changes the social identities we take up*.<sup>15</sup> It gets us to adopt new social identities and

13. The reference to *perceived* differences means that the categories that individuals bring to their understanding of the world, their motives, interests, etc, also matter to this kind of classification (Turner et al. 1994).

14. See Haslanger (forthcominga) for illuminating discussion of how social structures and agency relate.

15. The common ingroup identity model was put forward as a way of explaining Allport's contact hypoth-



to change the ranking of importance of our different social identities. In doing so, we become less motivated to single-mindedly defend pre-existing identities. Perhaps, as in the company merger case, those identities have disappeared from our self-concept, in which case we no longer need to defend them. Or, as in the university student case, perhaps we now need to balance out defending different social identities. Alternatively, having a richer self-concept, we find attacks to any one of its dimensions less threatening.

### 3.2 Interventions to Reshape Social Identities that Sustain Resistant Social Beliefs

Given the discussion above, re-shaping social networks can have deep effects on the structure of the self-concept. I will now consider two specific ways of re-shaping social networks that can get individuals to systematically reduce their attachment to social identities that support resistant social beliefs.

#### 3.2.1 Dispersed social networks

Consider a social network where individuals have a large number of diverse social ties, belonging to a number of different community spaces which they flexibly enter and exit. Someone who navigates very different social environments in their workplace, religious community, and their children's school district will have such a social network. Such a social network will often foster a more positive role for social identities in one's self-concept than a network where one belongs to a single, rigid, clearly delimited social group.<sup>16</sup>

In particular, such dispersed social networks provides many opportunities for cross-group contact. Correspondingly (given the common ingroup identity model), they make room for individuals to adopt new social identities. Having a self-concept that incorporates a rich diversity of social identities makes individuals less likely to single-mindedly defend any particular one. It generates the need to balance out defending different social identities and it reduces the overall threat posed by attacks to a specific dimension of the self.

Further, individuals will find different social identities (and conceptions thereof) salient as they move across these different social contexts. This will provide contexts for open-mindedness on specific topics. Perhaps at church, where religious identity is highly salient, a person is likely to defend traditionalist beliefs about gender roles; but, when talking with people at their mixed-gender running team, much less so.

---

esis (Allport 1954) that prejudice can be reduced by cross-group contact in good conditions. I am here repurposing the common ingroup identity model to think about addressing identity-protective reasoning (as opposed to explaining the modulation of affective dimensions of prejudice).

16. Such dispersed social networks need not amount to fully integrated societies, in the sense of "comprehensive intergroup association on terms of equality...[which] requires the full inclusion and participation as equals of members of all races in *all* social domains" (Anderson 2010, 112) (my italics). The existence of cross-group contact under positive conditions is compatible with the existence of community spaces which are not open to all. It only requires the existence of *some* spaces where boundaries are broken. Given concerns about the costs of integration for marginalized groups (e.g., loss of access to services that cater to their specific needs and of distinctive cultural and sense-making spaces (Shelby 2017)), I think we should be careful when proposing integration as the way to promote helpful cross-group contact.

Finally, having a social network that is mostly constituted of a large number of weak social ties, as opposed to a small number of strong ones, is likely to make social identities recede in importance in comparison with personal characteristics (Mutz 2006). This may reduce how much people engage in identity-protective reasoning across the board.

Focusing on partisan social identities makes this dynamic clearer. Most partisans in the US now live in partisan bubbles (Bishop 2009, Levendusky 2009). They barely ever interact with supporters of the other party. As a result, they have ceased to have cross-cutting social identities (i.e., social identities that encompass both Democrats and Republicans). As Democrats and Republicans share fewer and fewer (non-political) identities, they grow more invested in defending their partisan identities. Their entire sense of self stands and falls with their partisan identity.

If social sorting is the problem, cross-group contact might be the solution. There is evidence that Democrats who have cross-cutting affiliations, thereby sharing important identities with Republicans (and vice-versa) are less likely to engage in partisan identity-protective reasoning (Mason 2018). If we want to reduce partisan identity-protective reasoning, building networks that make room for shared identities among Democrats and Republicans is a good idea.<sup>17</sup>

### 3.2.2 Social movements

Social movements provide opportunities for contact for people of different social identities who share a commitment to the same goal, and for conversations that aim at building solidarity and community. Such conversations tend to result in our noticing commonalities and coming to identify under a common header. Because they function as sub-community spaces that foster solidarity, social movements are exemplary sites for the kind of cross-group contact that produces shared identities among members of distinct social groups.

Social movements offer additional benefits when it comes to identity-protective reasoning. A social movement is a “sustained campaign of claim-making, using repeated performances that advertise the claim, based on organizations, networks, traditions, and solidarities that sustain these activities.” (Tilly and Tarrow 2015, 11). Unlike weak social ties, participation in social movements tends to generate strong senses of affiliation and community. Further, social movements tend to generate novel strong social identities, and to lead participants to reshape their sense of self around new ones.

Social movements often creatively produce new social identities around their shared goals. This can happen as a by-product of collective action, as we witness with labor organizer identities. Through participation in the labor movement, racially inclusive union identities come to the fore and white identities recede. This leads union membership to reduce racial resentment among white workers and increase support for policies that benefit Black people (Frymer and Grumbach 2021).

Other times, social movements explicitly aim to produce or reshape social identities. For example, the Civil Rights movement explicitly aimed to produce a new

<sup>17</sup>. Networks with a large number of weak social ties have additional benefits for awareness of rationales for opposing views and for political tolerance and support for civil liberties (Mutz 2006).

Black American identity that did not incorporate white-enforced stereotypes. In Martin Luther King's words, one goal of the movement was to make room for "the new Negro," "a person with a new sense of dignity and destiny, with a new self-respect" (Martin Luther King 1956). In a different political direction, the NRA has had great political success largely by managing to purposefully cultivate "a distinct, politicized gun owner social identity" (Lacombe 2019, 1342).

Further, social movements sometimes attempt to produce alternative conceptions of identities seemingly unrelated movement participation. For example, the vegan movement has invested in building new images of masculinity that do not incorporate meat-eating. It has done so by bringing attention to hyper-masculine vegan athletes and highlighting how veganism can be taken to express certain stereotypically masculine traits, such as emotional stoicism and protecting others (Greenebaum and Dexter 2018).

In producing and inculcating social identities, participation in social movements often leads to changes in participants' self-concepts (Kiecolt 2000). Participants might discard or add new social identities to their self-concept. For example, they might cease to identify as victims and start identifying as activists. They might come to center some social identities more than they did before, and come to devalue identities that were once central to them. In this way, upper-middle-class women who participated in the feminist movement in the 70s often began to think of themselves as feminists, and to identify less as housewives (Breinlinger and Kelly 2014). Participants may also come to reconceptualize identities in ways that match those of their activist community, thereby changing their self-conception. A man who becomes involved in the vegan movement might not only come to center a vegan social identity in his self-conception, but also come to reconceptualize masculinity. Finally, if the social movement gets enough public attention, these identities have a chance to turn into culturally mainstream identities (or ways of conceptualizing a given identity), as has happened with feminist identities.<sup>18</sup>

In sum, involvement in social movements can powerfully reshape identities by leading participants to adopt or center (1) identities as members of that social movement, (2) other movement-relevant identities, or identities conceptualized as the social movement proposes, and (3) identities that are inclusive of people who were previously exclusively outgroup members.

This will re-orient identity-protective reasoning in important ways. Defending masculinity under a conception where it does not involve meat-eating is compatible with being open-minded on evidence about the health of meat-eating or the environmental effects of meat consumption. Ceasing to primarily defend whiteness, and instead defending one's identity as a union organizer, makes one more open-minded when it comes to racial injustice (including factual questions about American history or the racial distribution of wealth in America)—especially if this new identity is racially inclusive.

---

18. See Amenta and Polletta 2019 for an overview of the effects of social movements on broader culture.

### 3.3 Social Network Change as a Structural Intervention

I have suggested two ways of reshaping social networks that systematically restructure the self-concept. Given the role of the self-concept in generating resistance to counter-evidence, these interventions can address the problem of resistant social beliefs. Which kind of intervention is appropriate can only be decided on a case-by-base basis, and requires us to think of the broader effects of such interventions. In some cases, the best route to addressing resistant social beliefs will involve developing and coming to center strong alternative identities. In others, it will involve reducing attachment to social identities in general or make room for more inclusive social identities. Correspondingly, we might want to prioritize developing social movements (in the first case) or dispersed social networks (in the second case).

It is beyond the scope of this paper to develop concrete suggestions for such interventions. Instead, the discussion is meant to generate a blueprint for investigating (and testing) such interventions for specific resistant social beliefs.

The key point for my purposes is the following: social network change can reduce identity-protective reasoning, making individuals open to changing their minds on socially significant topics. In this way, social network change can play a key role in solving the problem of resistant social beliefs. And, importantly, social network change is a form of structural change: a large-scale change in the context in which beliefs are formed and maintained.

There is wide consensus in the literature in classifying social network change as a form of structural change. O'Connor and Weatherall (2019), Madva (2020) explicitly list it as a structural intervention that contrasts with individual-level debiasing, and Anderson (2010)'s integrationist proposals (a form of social network change) are generally taken to be structuralist.

One might object that changes in social networks also involve changes in individual attitudes. Individuals must come to enjoy spending time with a different set of people, and they might need to come to know their way about different neighbourhoods, learn to appreciate different ways of doing things, and so on. But this point applies to all sorts of reforms that we naturally describe as structural. For example, legal changes, or changes in the material set-up which individuals navigate, require specific agents to implement them, and, as such, involve individual attitudes.

To the extent that it is helpful to describe some interventions as structural and others as individual-level, changes in social networks are best construed as structural changes.<sup>19</sup> Specifically, they are not well-modeled by thinking of isolated individuals changing their minds one-by-one. Instead, they require large-scale and highly coordinated changes in practices, social norms, the material conditions in which individuals interact, and attitudes that sustain current social networks. As is characteristic of structural changes, non-coordinated individual action will have little impact (Haslanger forthcomingb).

---

19. Brownstein et al. (2021) argue against this distinction, proposing that it is a mere matter of framing. Of course, if you reject the distinction between practical individualism and structuralism, the question of the connection between methodological and practical positions in the individualism vs. structuralism debate does not arise. If you agree with Brownstein et al. (2021), the key take-away from this paper is a concrete, empirically-informed proposal for a family of interventions that can address identity-protective reasoning.

An analogy will help see the structural nature of the changes I am suggesting. Consider the case of addressing psychiatric disorders. An individualist approach would prescribe a combination of medication, talk therapy, therapeutic “homework” exercises, and the like. It would not directly aim to change the backdrop of the person’s life, that is, their economic, material, and social conditions. In contrast, a structuralist approach would note the ways in which psychiatric disorders can be sustained by lack of life prospects, economic difficulties, social marginalization and isolation, among other structural factors, and aim to address such problems as a way of addressing such disorders.<sup>20</sup>

The same distinction translates over to the non-pathological case of identity-protective reasoning. Individualist interventions might propose information workshops or training sessions to help individuals avoid identity-protective reasoning (Madva 2017). These are analogues of therapy or medicating in psychiatric or therapeutic contexts. These interventions contrast with changes in the social networks which individuals navigate day-to-day in their social interactions. By changing such background conditions to our day-to-day life, we can address resistant social beliefs—much like we might be able to address different forms of psychological suffering by changing living conditions.

The central upshot of this section, then, is the following: against initial appearances, a form of structural change is a powerful lever for addressing identity-protective reasoning. Consequently, the practical structuralist has resources for changing even resistant, active beliefs that are protected against counter-evidence—the kinds of beliefs that have been claimed to be intractable for the practical structuralist.

## 4 The Deep Effects of Social Network Change on Psychology

Identity-protective reasoning is not the only factor that sustains resistant social beliefs. Fully addressing the problem of resistant social beliefs will require the practical structuralist to have resources to address other sources of the persistence of these beliefs.

Though it is beyond the scope of a single paper to consider how to address *all* sources of resistant social beliefs, I want to bolster the case for the role of structural interventions by showing that they can powerfully influence many central aspects of belief revision. Correspondingly, in this section, I will argue that social network change can also address other crucial factors that help maintain resistant social beliefs.

### 4.1 Other kinds of motivated reasoning

According to systems justification theory (Jost 2019), desires to justify the social *status quo* can influence our reasoning (much like identity-protective desires do). These

---

20. See Pickard 2020 for discussion of this distinction in the context of addiction treatment.

desires arise because feeling good about the *status quo* increases satisfaction and reduces the uncertainty, threat, and social discord that would come from attempting to bring about social change (Jost 2019). And, much like identity-protective desires, these desires have long-ranging effects. They protect both beliefs that the current social system is good and beliefs that serve to justify it or make it appear natural (e.g. beliefs in the natural submissiveness of women or in the supposed meritocratic nature of our society). More generally, desires to justify the social *status quo* help maintain ideological beliefs (Haslanger 2011, Haslanger 2017, Shelby 2003).

Social movements, it turns out, can also address such desires. Given the source of these desires, reducing the felt uncertainty, threat, and social discord that would come from attempting to bring about social change, and providing sources of satisfaction that are compatible with disliking the *status quo*, will help reduce systems justification reasoning. Social movements can play a powerful role in both of these. In building community and making space for experiments in living (Anderson 2014), social movements provide new sources of satisfaction that are compatible with disliking the *status quo*. Indeed, given the oppositional nature of social movements, accessing these sources of satisfaction might actively involve negative attitudes towards the reigning order. Further, as a form of collective action, social movements make attempting to bring about social change less daunting (Haslanger forthcomingb). In these ways, they reduce the costs of questioning the *status quo* for individuals.

A second relevant set of desires behind resistant social beliefs are desires to have socially adaptive beliefs (Williams 2021), i.e., desires to have beliefs that we are socially rewarded for having and to avoid beliefs that we would be socially punished for having. Socially adaptive beliefs include beliefs that serve to signal allegiance to the groups to which we belong (e.g., beliefs about gun use when it comes to partisanship (Kahan 2012)); beliefs that are required for good standing in a group (e.g. beliefs about the literal truth of the Bible for many mainstream Christian groups); and beliefs that facilitate smooth participation in dominant cultural practices (e.g., beliefs in the beauty of small feet, which promoted participation in foot-binding in early 20th-century China (Mackie 1996, Sankaran 2020)). The desire to have socially normative beliefs leads people to interact with evidence in ways that promote maintaining those beliefs.<sup>21</sup>

The kinds of social network change that I discussed in §3.2 can reshape the desires that sustain socially adaptive beliefs. It is a familiar point that social movements affect social norms (e.g., Anderson 2014, Bicchieri 2016, Haslanger 2015, Haslanger 2017, Haslanger 2019 Sankaran 2020). Different sub-communities develop and enforce different social norms, including social norms on belief. To the extent that individuals are immersed in different communities, they become responsive to different sets of social norms. In the context of the social movement, norms that pressured them to have a given belief in dominant contexts do not have a grip on them, making them open to counter-evidence.

---

21. There is some disagreement about whether such cases reflect real beliefs. Some theorists (Hannon 2021, Schaffner and Luks 2018) think that social pressure only motivates people to *appear* to have certain beliefs, which can be done without actual belief. For reasons that Quilty-Dunn (2020) and Williams (2021) point out, I think that at least some such cases amount to belief. If you disagree, you can just bracket this discussion, as the central point of the paper does not depend on it.

Social movements also often play a key role in changing social norms writ large: a vocal minority opposing a norm can lead to wide reconsideration of the norm, and to the norm losing its authority (Anderson 2014). For example, public pledges by sub-communities played an important role in changing social norms about foot-binding in early 20th century China (Mackie 1996), including norms on *beliefs* about foot-binding. Such public pledges reduce the social costs of not following previously unquestioned norms. For example, the social costs to rejecting the (previously dominant) beliefs that foot-binding promotes good health and fertility were reduced as more and more people came to vocally reject such views.

Indeed, social movements can affect beliefs without any new evidence coming into the picture. In developing and disseminating new social norms, social movements might make new beliefs socially normative, generating desires to have and maintain those new beliefs. For example, as campaigns against foot-binding succeeded, the belief that foot-binding is cruel and bad for one's health became socially normative, generating pressure for individuals to have this belief. Similarly, as we saw at the end of §3, social movements can change which social identities we defend, making us engage in identity-protective reasoning with respect to a different set of beliefs.<sup>22</sup>

In this way, belief change can be entirely driven by changes in social networks and norms and the changes in desires that these engender. This is an old thought: as Pascal (1852) put it, if you want to adopt certain beliefs, you should “Endeavour to convince yourself, not by increase of proofs, but by the abatement of your passions,” where a crucial means to this is integrating yourself in communities that have the relevant beliefs. My discussion puts flesh to the bones of this point, by drawing on research in psychology to explain why Pascal's suggestion can succeed. Specifically, changing social communities changes many of our desires, which have to do with fitting in, having positive relationships with those closest to us, and feeling positive about those we see as ‘our people’. Given that such desires profoundly affect belief revision, changing social communities can entirely re-orient our doxastic take on the world in a way that bypasses evidence.

Alternatively, social network change can genuinely make agents more open to evidence—not merely bump them from one belief to another. Much as social network change can make us engage less in pernicious forms of identity-protective reasoning, it can destabilize what beliefs count as normative without replacing them with new normative beliefs in the same domain. In this way, social network change can set the stage for agents to rationally respond to evidence, and thereby for genuine rational engagement.<sup>23</sup>

---

22. I am not recommending this as a strategy for changing others' beliefs. In bypassing agents' epistemic rationality, this would be problematically manipulative. Deliberately implementing such a strategy risks eroding important democratic ideals of mutual respect and collective deliberation. My point is simply that changes in social networks can in and of themselves have deep effects on our beliefs. Insofar as we are interested in either understanding or promoting belief change in a social context, we need to be alert to this phenomenon.

23. Here too one might have ethical worries about policy-makers devising and implementing social network change as a means to collective epistemic improvement, even if evidence and argument come in. Whether and when such interventions would infringe on our autonomy, and the limits of their moral permissibility, are open questions, and ones which would require engaging with the details of specific interventions.

## 4.2 Trust and access to evidence

Suppose that motivational factors have been cleared away and agents are open to evidence. This on its own does not suffice to get agents to abandon their resistant social beliefs. They must, additionally, have access to evidence that challenges those beliefs, and count it as evidence in the first place. For this reason, solving the resistant social beliefs problem requires good distribution of evidence and appropriate patterns of trust (of the kind that allow us to properly count others' good testimony as evidence).

Here, again, the shape of social networks turns out to be crucial. As recent work in social epistemology has emphasized, the fact that agents sometimes have biased samples of evidence and fail to trust reliable sources often has its roots in dysfunctional social networks. O'Connor and Weatherall (2019) persuasively argue that the shape of social networks affects the evidence agents have in ways that help explain the maintenance of false beliefs. Nguyen (2020) argues that *echo chambers*, i.e. social epistemic structures which pervert and corrupt one's epistemic trust, pose additional problems that are independent of evidence access.

Restructuring social networks has a crucial role to play in addressing such problems. Opportunities for cross-group contact provide occasions for epistemic friction (Medina 2013), including for receiving counter-arguments to one's views. Such opportunities can also make individuals open to a wider range of sources of evidence. Group belonging affects trust. We tend to trust ingroup members and distrust outgroup members (Tajfel 1970), going as far as rejecting information offered by outgroup members while accepting the same information from ingroup members (Turner et al. 1987). To the extent that cross-group contact reshapes who counts as being in one's ingroup, they will reorient one's trust.<sup>24</sup>

Reshaping social networks can also improve one's patterns of trust by reducing testimonial injustice. Testimonial injustice occurs when a hearer's identity prejudice leads them to ascribe less credibility than the speaker deserves (Fricker 2007). Reducing identity prejudice helps address testimonial injustice. To the extent that identity prejudice is constituted by prejudiced beliefs, and prejudiced beliefs are partly maintained in virtue of identity-protective reasoning, the interventions outlined in §3 will contribute to addressing testimonial injustice.<sup>25</sup>

In addition to changing evidence access and patterns of trust, reshaping social networks can affect agents' perspectives (Camp 2017). As Fraser (2021) has compellingly argued, testimonial exchanges often do more than transmit evidence. They convey perspectives, i.e. suites of interlocking dispositions to attend, inquire, value, and interpret the world in specific ways (Camp 2017). To the extent that changing social networks changes whose testimony individuals hear, it expands our access to dif-

24. This observation does not provide a full solution to the echo chambers problem: as Nguyen (2020) notes, it will be very difficult to get people to interact with outgroup members and to feel a sense of kinship with them if their starting point is demonizing outgroup members. The point is simply that, if we could get such interactions going, we should expect to see re-orientations of trust.

25. Fully establishing the role of social network change in addressing testimonial injustice would require determining to what extent identity prejudice is implemented in prejudiced beliefs, and how large the role of identity-protective reasoning is in the maintenance of prejudiced beliefs. If this line of reasoning succeeds, it offers further support structural interventions (Anderson 2012) as opposed to virtue cultivation (Fricker 2007, Madva 2019) to address testimonial injustice.



ferent perspectives. For this reason, cross-group contact can help us develop what [Arendt \(1989\)](#) calls an “enlarged mentality”, characterized by having a wide range of standpoints present in one’s mind and the imaginative capacity to occupy them.

In sum, social networks affect a wide range of *psychological* factors involved in the maintenance of resistant social beliefs. Given the importance of social networks in shaping our belief maintenance and revision—by affecting evidence access, trust, perspectives, and patterns of motivated reasoning—they should be at the front and center of our theorizing about socially significant belief change. This has implications for how we think of the interaction between structural and individual change. I turn to these in the next section.

## 5 Integrating Methodological Individualism and Practical Structuralism

In the introduction, I noted that theorists of social change tend to split into two polarized camps: individualists and structuralists. Each side takes a distinctive party line on both methodology and practical recommendations. Individualists recommend detailed study of psychology and interventions that focus on changing hearts and minds one-by-one. At the other end of the spectrum, structuralists eschew the study of the mind and propose changes to large-scale social structures.

The discussion in this paper challenges this polarized field. We can and should integrate a measure of methodological individualism into our study of social change. Doing so is compatible with, and may even support, practical structuralism.<sup>26</sup>

Let’s start with practical structuralism. In this paper, I have argued that change in social networks is a powerful means for achieving change in individual beliefs. Social network change, I have argued in §3.3, is a form of structural change. Therefore structural reform can be a powerful lever for changes in individual belief.

This is a novel point for practical structuralism. Structuralists have emphasized that some components of unjust social structures persist independently of the beliefs of (the vast majority of) individuals in that structure. Think here of the material set-up of buildings that are not accessible to people with physical disabilities or of complex bureaucracies. We need structural interventions to target such components. But, as we saw in §1, we also need interventions targeting beliefs. Individualist-leaning theorists have taken this to checkmate structuralism. For instance, [Madva \(2020\)](#) writes that cases where our beliefs push back against structural change provide “an argument for insisting on the importance of *individual level* debiasing strategies, which change individual’s biased assumptions” ([Madva 2020](#), my italics).

Given my discussion in §3 and §4, this conclusion is unwarranted. Structural interventions can play a key role in changing beliefs. To put it differently: prioritizing belief change is compatible, and may even support, prioritizing structural interventions.<sup>27</sup> This paper shows that structural interventions can powerfully target even

26. Though this combination of views has been neglected in recent literature, it harks back to [Fanon \(2007\)](#)’s discussion of social change, which combines centering deep structural reform with attention to the deep psychological effects of oppression.

27. Two qualifications. First, accepting the power of structural interventions for belief change is compatible

resistant social beliefs—the cases that are supposed to pose the deepest problem for structuralism. In doing so, it substantially strengthens the case for practical structuralism.

I will turn, now, to considering methodological aspects of the structuralism–individualism debate. My discussion above pays close attention to the psychology of belief maintenance. This goes against recent trends in social epistemology, which explicitly reject appealing to psychological factors in explaining socially troubling beliefs. For example, O’Connor and Weatherall (2019) write that “to focus on individual psychology is to badly misdiagnose how false beliefs persist and spread” (O’Connor and Weatherall 2019, 7). Similarly, Thi Nguyen writes that we should not understand the persistence of socially pernicious beliefs “in terms of individual psychological tendencies, such as motivated reasoning” but of “systems and environments” (Nguyen 2021, 231).

As I discussed in §4, these projects yield many important insights about the role of social networks in shaping evidence access and trust. However, eschewing appeal to psychological factors is misguided. As I argued in §2, we have strong empirical reasons to think that identity-protective reasoning plays a role in the maintenance of resistant social beliefs. Ignoring its role leads to theoretical distortions in our understanding of the mind and of social change. It also leads to unpleasant practical surprises. Interventions that ignore identity-protective reasoning will encounter active resistance from believers. Without changes in desires and motivation, counter-evidence (from trusted sources) is likely to be resisted. For instance, when people are motivated to maintain a stereotypical belief, counter-evidence to that belief tends to result in sub-typing (i.e., coming to believe that the stereotype applies to a subset of the original group), not in the abandonment of the stereotype (Kunda and Oleson 1995, Richards and Hewstone 2001).

Further, against some structuralists’ insistence on the irrelevance of psychology, *any* form of theorizing about belief change involves making assumptions about the structure of cognition. Despite their claim to eschew psychology, both O’Connor and Weatherall (2019) and Nguyen (2021) make assumptions about the psychology of individual agents: O’Connor and Weatherall (2019) assume a Bayesian cognitive architecture, and Nguyen (2021) assumes bounded rationality (i.e. ideal (Bayesian) rationality bounded by processing limitations).

The charitable way of reading these projects is as making use of idealization: they idealize away substantive deviations from simple rational models of cognition. Such idealization helpfully isolates the role of factors beyond such deviations in belief maintenance. However, it is important to keep in clear sight that we are idealizing, and to recognize that a complete explanation will have to bring in accurate psychological models.

Indeed, to the extent that structuralists rely exclusively on idealized models of cognition, they make themselves vulnerable to worries that “the policy predictions

---

with also incorporating individualist interventions, such as individual debiasing (Madva 2017). Second, I am presupposing, as structuralists generally do, that there is a worthwhile distinction to be drawn between structural and individual interventions, even if structural interventions typically require individual action (Ayala-López and Beeghly 2020) and whether we are inclined to classify an intervention as structural is subject to framing effects (Brownstein et al. 2021).

of structural prioritizers rely on oversimplified psychological models” (Madva 2016, 702). Specifically, Madva thinks that structuralism assumes a picture in which we get individual change for free from independently desirable structural reforms: *The Mirroring View of Beliefs*, i.e. the view that beliefs are “mirror-like reflections of local environments and communities within which individuals are immersed” (Dasgupta 2013, 240), reflecting the bad evidence that individuals have available to them. As Madva correctly points out, this view is false: individual minds actively resist counter-evidence in ways that allow us to maintain socially troubling beliefs.

It is true that structuralists have often endorsed something like the Mirroring View (e.g. Antony 2016, Huebner 2016). My discussion in this paper shows that practical structuralism does not rely on this view. Focusing on the ways in which we actively resist the evidence in fact *supports* structural interventions. Indeed, structuralists have independent reasons to be suspicious of the Mirroring View. Structuralists often suggest that our cognitive structures are partially the result of internalizing social structures (Zheng 2018). Internalizing social structures involves actively filtering the world in ways that express the effect of social norms and identities on our epistemic agency. It is open to the structuralist to think of the structural context for belief revision as including aspects of our minds that are deeply shaped by social structures, such as identity-protective reasoning. For these reasons, structuralists should be much friendlier to psychology than they currently tend to be.

## 6 Conclusion

In theorizing about social change, paying close attention to the psychology of belief change, and especially to “deviant” factors such as identity-protective reasoning, is typically taken to lead to all-encompassing individualism about social change. Against this, I have argued that structural change (in the form of social network change) can address identity-protective reasoning. Indeed, changing the shape of social networks can deeply restructure our epistemic relationship with the world, affecting the evidence we have available, who we trust, and a range of forms of motivated reasoning.

Altogether, this brings a new perspective to the debate between individualism and structuralism. Individualism (and structuralism) about methodology and interventions are separable. Methodological individualism need not support individualist interventions. Focusing on psychology can *support*, instead of undermining, structural interventions, and it is compatible with recognizing the deep effects of social structures on psychological structures.

Noticing these points makes room for integrating important insights from both individualist and structuralist traditions. As structuralists suggest, it is important to attend to individuals *qua* nodes in social structures, and recognize how deeply structures shape individual minds. But, as individualists emphasize, this requires integrating realistic models of human psychology. In doing so, we can recognize the power of structural change for changing beliefs without seeing beliefs as simple mirrors of social reality.

## Works Cited

- Linda Martín Alcoff. Epistemologies of ignorance: Three types. In Shannon Sullivan Nancy Tuana, editor, *Race and Epistemologies of Ignorance*. SUNY Press, 2007.
- Michelle Alexander. *The New Jim Crow: Mass Incarceration in the Age of Colorblindness*. The New Press, 2010.
- Gordon Allport. *The Nature of Prejudice*. Addison-Wesley, 1954.
- Edwin Amenta and Francesca Polletta. The cultural impacts of social movements. *Annual Review of Sociology*, 45:279–299, 2019.
- Elizabeth Anderson. *The imperative of integration*. Princeton University Press, 2010.
- Elizabeth Anderson. Epistemic justice as a virtue of social institutions. *Social epistemology*, 26(2):163–173, 2012.
- Elizabeth Anderson. Social movements, experiments in living, and moral progress: Case studies from Britain’s abolition of slavery, 2014.
- Louise Antony. Bias: Friend or foe? Reflections on Saulish skepticism. *Implicit bias and philosophy*, 1:157–190, 2016.
- Hannah Arendt. *Lectures on Kant’s political philosophy*. University of Chicago Press, 1989.
- Beth Avery and Han Lu. Ban the box: U.s. cities, counties, and states adopt fair hiring policies, Oct 2021. URL <https://www.nelp.org/publication/ban-the-box-fair-chance-hiring-state-and-local-guide/>.
- Saray Ayala-López and Erin Beeghly. Explaining injustice: Structural analysis, bias, and individuals. In *An Introduction to Implicit Bias*, pages 211–232. Routledge, 2020.
- Cristina Bicchieri. *Norms in the wild: How to diagnose, measure, and change social norms*. Oxford University Press, 2016.
- Bill Bishop. *The big sort: Why the clustering of like-minded America is tearing us apart*. Houghton Mifflin Harcourt, 2009.
- Sara Breinlinger and Caroline Kelly. *The social psychology of collective action*. Taylor & Francis, 2014.
- Rupert Brown. Social identity theory: Past achievements, current problems and future challenges. *European Journal of Social Psychology*, 30(6):745–778, 2000.
- Michael Brownstein, Daniel Kelly, and Alex Madva. Individualism, structuralism, and climate change. *Environmental Communication*, pages 1–20, 2021.
- Elisabeth Camp. Perspectives in imaginative engagement with fiction. *Philosophical Perspectives*, 31(1):73–102, 2017.

- Nilanjana Dasgupta. Implicit attitudes and beliefs adapt to situations: A decade of research on the malleability of implicit prejudice, stereotypes, and the self-concept. *Advances in experimental social psychology*, 47:233–279, 2013.
- Lacey J Davidson and Daniel Kelly. Minding the gap: bias, soft structures, and the double life of social norms. *Journal of Applied Philosophy*, 37(2):190–210, 2020.
- Daniel Dellaposta. Pluralistic collapse: The “oil spill” model of mass opinion polarization. *American Sociological Review*, 85(3):507–536, 2020. doi: [10.1177/0003122420922989](https://doi.org/10.1177/0003122420922989).
- John Dixon, Mark Levine, Steve Reicher, and Kevin Durrheim. Beyond prejudice: Are negative evaluations the problem and is getting us to like one another more the solution? *Behavioral and Brain Sciences*, 35(6):411–425, 2012.
- Jennifer L Doleac and Benjamin Hansen. Does “Ban the Box” help or hurt low-skilled workers? statistical discrimination and employment outcomes when criminal histories are hidden. Technical report, National Bureau of Economic Research, 2016.
- Tessa L Dover, Brenda Major, and Cheryl R Kaiser. Diversity initiatives, status, and system-justifying beliefs: When and how diversity efforts de-legitimize discrimination claims. *Group Processes & Intergroup Relations*, 17(4):485–493, 2014.
- John F Dovidio, Alice M Isen, Paula Guerra, Samuel L Gaertner, and Mary Rust. Positive affect, cognition, and the reduction of intergroup bias. In *Intergroup cognition and intergroup behavior*, pages 337–366. Lawrence Erlbaum Associates Publishers, 1998.
- Frantz Fanon. *The wretched of the earth*. Grove/Atlantic, Inc., 2007.
- Rachel Fraser. Narrative testimony. *Philosophical Studies*, pages 1–28, 2021.
- Miranda Fricker. *Epistemic injustice: Power and the ethics of knowing*. Oxford University Press, 2007.
- Marilyn Frye. *The politics of reality: Essays in feminist theory*. Crossing Press, 1983.
- Paul Frymer and Jacob M Grumbach. Labor unions and white racial politics. *American Journal of Political Science*, 65(1):225–240, 2021.
- Samuel L Gaertner, John F Dovidio, Jason A Nier, Brenda S Banker, Christine M Ward, Melissa Houlette, and Stephenie Loux. The common ingroup identity model for reducing intergroup bias: Progress and challenges. In *Social identity processes: Trends in theory and research*, pages 133–148. Sage Publications Ltd, 2000.
- Jorge LA Garcia. The heart of racism. *Journal of social philosophy*, 27(1):5–46, 1996.
- Steffen R Giessner, Johannes Ullrich, and Rolf van Dick. A social identity analysis of mergers & acquisitions. *The handbook of mergers and acquisitions*, pages 474–494, 2012.

- Daniel Gilbert. *Stumbling on happiness*. Vintage Canada, 2009.
- Jessica Greenebaum and Brandon Dexter. Vegan men and hybrid masculinity. *Journal of Gender Studies*, 27(6):637–648, 2018.
- Michael Hannon. Disagreement or badmouthing? the role of expressive discourse in politics. In Michael Hannon and Elizabeth Edenberg, editors, *Political Epistemology*. Oxford University Press, 2021.
- Sally Haslanger. Ideology, generics, and common ground. In *Feminist metaphysics*, pages 179–207. Springer, 2011.
- Sally Haslanger. Distinguished lecture: Social structure, narrative and explanation. *Canadian Journal of Philosophy*, 45(1):1–15, 2015.
- Sally Haslanger. Racism, ideology, and social movements. *Res Philosophica*, 94(1):1–22, 2017.
- Sally Haslanger. Cognition as a social skill. *Australasian Philosophical Review*, 3(1): 5–25, 2019.
- Sally Haslanger. Failures of methodological individualism: the materiality of social systems. *Journal of Social Philosophy*, 2020.
- Sally Haslanger. Ideology in practice: What does ideology do? *The Aquinas Lecture*, forthcominga.
- Sally Haslanger. How to change a social structure. In Ruth Chang and Amia Srinivasan, editors, *Normative Philosophy: Conversations in Moral, Legal, and Political Philosophy*. Oxford University Press, forthcomingb.
- Michael A Hogg, John C Turner, and Barbara Davidson. Polarized norms and social frames of reference: A test of the self-categorization theory of group polarization. *Basic and Applied Social Psychology*, 11(1):77–100, 1990.
- Bryce Huebner. Implicit bias, reinforcement learning, and scaffolded moral cognition. In Michael Brownstein and Jennifer Saul, editors, *Implicit Bias and Philosophy, Volume 1: Metaphysics and Epistemology*. Oxford University Press, 2016.
- Matthew W Hughey. White backlash in the ‘post-racial’ united states. *Ethnic and Racial Studies*, 37(5):721–730, 2014.
- John T Jost. A quarter century of system justification theory: Questions, answers, criticisms, and societal applications. *British Journal of Social Psychology*, 58(2):263–314, 2019.
- Dan M Kahan. Ideology, motivated reasoning, and cognitive reflection: An experimental study. *Judgment and Decision making*, 8:407–24, 2012.
- Dan M Kahan. The expressive rationality of inaccurate perceptions. *Behavioral & Brain Sciences*, 40:26–28, 2015.

- Dan M Kahan. Misconceptions, misinformation, and the logic of identity-protective cognition. *SSRN Electronic Journal*, 2017.
- Cheryl R Kaiser, Brenda Major, Ines Jurcevic, Tessa L Dover, Laura M Brady, and Jennessa R Shapiro. Presumed fair: Ironic effects of organizational diversity structures. *Journal of personality and social psychology*, 104(3):504, 2013.
- K Jill Kiecolt. Self-change in social movements. *Self, identity, and social movements*, pages 110–131, 2000.
- William M Klein and Ziva Kunda. Motivated person perception: Constructing justifications for desired beliefs. *Journal of experimental social psychology*, 28(2):145–168, 1992.
- Joshua Knobe, Sandeep Prasada, and George E Newman. Dual character concepts and the normative dimension of conceptual representation. *Cognition*, 127(2):242–257, 2013.
- Ziva Kunda. The case for motivated reasoning. *Psychological bulletin*, 108(3):480, 1990.
- Ziva Kunda and Kathryn C Oleson. Maintaining stereotypes in the face of disconfirmation: Constructing grounds for subtyping deviants. *Journal of personality and social psychology*, 68(4):565, 1995.
- Matthew J Lacombe. The political weaponization of gun owners: The national rifle association’s cultivation, dissemination, and use of a group social identity. *The Journal of Politics*, 81(4):1342–1356, 2019.
- Matthew Levendusky. *The partisan sort*. University of Chicago Press, 2009.
- Akiva Liberman and Shelly Chaiken. Defensive processing of personally relevant health messages. *Personality and Social Psychology Bulletin*, 18(6):669–679, 1992.
- Gerry Mackie. Ending footbinding and infibulation: A convention account. *American sociological review*, pages 999–1017, 1996.
- Alex Madva. A plea for anti-anti-individualism: How oversimple psychology misleads social policy. *Ergo, an Open Access Journal of Philosophy*, 3, 2016.
- Alex Madva. Biased against debiasing: On the role of (institutionally sponsored) self-transformation in the struggle against prejudice. *Ergo: An Open Access Journal of Philosophy*, 4, 2017.
- Alex Madva. The inevitability of aiming for virtue. In Benjamin R. Sherman and Stacey Goguen, editors, *Overcoming Epistemic Injustice*, pages 85–99. Rowman and Littlefield International, 2019.
- Alex Madva. Individual and structural interventions. In *An Introduction to Implicit Bias*, pages 233–270. Routledge, 2020.

- Eric Mandelbaum. Attitude, inference, association: On the propositional structure of implicit bias. *Noûs*, 50(3):629–658, 2016.
- Eric Mandelbaum. Troubles with Bayesianism: An introduction to the psychological immune system. *Mind & Language*, 34(2):141–157, 2019.
- Jr. Martin Luther King. “the ‘New Negro’ of the South: Behind the montgomery story. *Socialist Call*, 24:16–19, 1956.
- Lilliana Mason. *Uncivil agreement: How politics became our identity*. University of Chicago Press, 2018.
- José Medina. *The epistemology of resistance: Gender and racial oppression, epistemic injustice, and the social imagination*. Oxford University Press, 2013.
- Charles Mills. White ignorance. In Shannon Sullivan Nancy Tuana, editor, *Race and epistemologies of ignorance*, pages 26–31. SUNY Press, 2007.
- Diana C Mutz. *Hearing the other side: Deliberative versus participatory democracy*. Cambridge University Press, 2006.
- C Thi Nguyen. Echo chambers and epistemic bubbles. *Episteme*, 17(2):141–161, 2020.
- C Thi Nguyen. The seductions of clarity. *Royal Institute of Philosophy Supplements*, 89: 227–255, 2021.
- Cailin O’Connor and James Owen Weatherall. *The misinformation age*. Yale University Press, 2019.
- Blaise Pascal. *Pensées*. Dezobry et E. Magdeleine, 1852.
- Jared Piazza, Matthew B Ruby, Steve Loughnan, Mischel Luong, Juliana Kulik, Hanne M Watkins, and Mirra Seigerman. Rationalizing meat consumption. the 4 Ns. *Appetite*, 91:114–128, 2015.
- Hanna Pickard. Addiction and the self. *Noûs*, 2020. doi: <https://doi.org/10.1111/nous.12328>.
- Tom Pyszczynski, Sheldon Solomon, and Jeff Greenberg. Thirty years of terror management theory: From genesis to revelation. In *Advances in Experimental Social Psychology*, volume 52, pages 1–70. Elsevier, 2015.
- Jake Quilty-Dunn. Unconscious rationalization, or: How not to think about awfulness and death. Manuscript, 2020.
- Zoeë Richards and Miles Hewstone. Subtyping and subgrouping: Processes for the prevention and promotion of stereotype change. *Personality and Social Psychology Review*, 5(1):52–73, 2001.
- Hank Rothgerber. Real men don’t eat (vegetable) quiche: Masculinity and the justification of meat consumption. *Psychology of Men & Masculinity*, 14(4):363, 2013.



- Bastiaan T Rutjens, Robbie M Sutton, and Romy van der Lee. Not all skepticism is equal: Exploring the ideological antecedents of science acceptance and rejection. *Personality and Social Psychology Bulletin*, 44(3):384–405, 2018.
- Kirun Sankaran. What’s new in the new ideology critique? *Philosophical Studies*, 177(5):1441–1462, 2020.
- Brian F Schaffner and Samantha Luks. Misinformation or expressive responding? what an inauguration crowd can tell us about the source of political misinformation in surveys. *Public Opinion Quarterly*, 82(1):135–147, 2018.
- Tommie Shelby. Ideology, racism, and critical social theory. *The Philosophical Forum*, 34(2):153–188, 2003.
- Tommie Shelby. *Dark ghettos*. Harvard University Press, 2017.
- David K Sherman and Geoffrey L Cohen. The psychology of self-defense: Self-affirmation theory. *Advances in experimental social psychology*, 38:183–242, 2006.
- Amia Srinivasan. Radical externalism. *Philosophical Review*, 129(3):395–431, 2020.
- Jason Stanley. *How propaganda works*. Princeton University Press, 2015.
- Claude M Steele. The psychology of self-affirmation: Sustaining the integrity of the self. *Advances in experimental social psychology*, 21(2):261–302, 1988.
- Shannon Sullivan. *Revealing whiteness: The unconscious habits of racial privilege*. Indiana University Press, 2006.
- William B Swann Jr. Seeking “truth,” finding despair: Some unhappy consequences of a negative self-concept. *Current Directions in Psychological Science*, 1(1):15–18, 1992.
- Olúfemi Táíwò. Beware of schools bearing gifts. *Public Affairs Quarterly*, 31(1):1–18, 2017.
- Henri Tajfel. Experiments in intergroup discrimination. *Scientific american*, 223(5):96–103, 1970.
- Henri Tajfel. Social psychology of intergroup relations. *Annual review of psychology*, 33(1):1–39, 1982.
- Henri Tajfel, John C Turner, William G Austin, and Stephen Worchel. An integrative theory of intergroup conflict. *Organizational identity: A reader*, 56(65):9780203505984–16, 1979.
- Elisha M Tarlow and David AF Haaga. Negative self-concept: Specificity to depressive symptoms and relation to positive and negative affectivity. *Journal of Research in Personality*, 30(1):120–127, 1996.
- Charles Tilly and Sidney G Tarrow. *Contentious politics*. Oxford University Press, 2015.

- John C Turner. Social categorization and the self-concept: A social cognitive theory of group behavior. In T. Postmes and N. R. Branscombe, editors, *Rediscovering social identity*, pages 243–272. Psychology Press, 2010.
- John C Turner and Penelope J Oakes. The significance of the social identity concept for social psychology with reference to individualism, interactionism and social influence. *British Journal of Social Psychology*, 25(3):237–252, 1986.
- John C Turner and Katherine J Reynolds. Self-categorization theory. *Handbook of theories in social psychology*, 2(1):399–417, 2011.
- John C Turner, Michael A Hogg, Penelope J Oakes, Stephen D Reicher, and Margaret S Wetherell. *Rediscovering the social group: A self-categorization theory*. Basil Blackwell, 1987.
- John C Turner, Penelope J Oakes, S Alexander Haslam, and Craig McGarty. Self and collective: Cognition and social context. *Personality and social psychology bulletin*, 20(5):454–463, 1994.
- Jay J Van Bavel and Andrea Pereira. The partisan brain: An identity-based model of political belief. *Trends in cognitive sciences*, 22(3):213–224, 2018.
- Daniel Williams. Socially adaptive belief. *Mind & Language*, 36(3):333–354, 2021.
- Robin Zheng. Bias, structure, and injustice: A reply to haslanger. *Feminist Philosophy Quarterly*, 4(1), 2018.