Structural Dimensions of Discrimination

Katharina Berndt Rasmussen
Stockholm University & Institute for Futures Studies
katharina.berndt@philosophy.su.se

Early draft, please do not cite or circulate.

Abstract

In this paper, I put forth a theory of discrimination that captures two structural dimensions of the phenomenon that are often pushed to the periphery of analysis. I outline a concept of discrimination which aims to capture what many of us are concerned with in real life cases, and explore its agential and structural forms, as well as its relation to moral wrongness and to social injustice. I propose that we should see discrimination as a bridge concept: between the personal and the political, between criteria of moral wrongness and of social (in)justice, between individualist and structural analyses of harms and inequalities. I believe that it is this bridging feature which makes the concept distinctive and particularly useful, both within philosophy and more broadly in society (politics, law, civil society).

I proceed as follows: in section 2, I propose a definition of discrimination and show how it gives rise to four distinct forms of discrimination. In section 3, I describe cases of implicit bias discrimination and epistemic injustice and argue that my definition can capture these, given a structural reading. This captures the first structural dimension of discrimination. In section 4, I give an account of the moral wrongness of discrimination, in terms of harm. I propose that this account still does not exhaustively capture what makes the phenomenon of discrimination intuitively problematic. I therefore, in section 5, explore the connection between discrimination and theories of social (in)justice. This captures the second, much more foundational, structural dimension of discrimination. Section 6 concludes.

1. Introduction

In this paper, I put forth a theory of discrimination that captures two structural dimensions of the phenomenon, which are often pushed to the periphery of analysis. I do this by outlining a concept of discrimination, which aims to capture what many of us are concerned with in real life cases, and by exploring the different forms it can take, as well as its relation to moral wrongness and to social injustice. I propose that we should see discrimination, in the relevant sense, as a bridge concept: between the personal and the political, between criteria of moral wrongness and of social (in)justice, between, indeed, individualist and structural analyses of harms and inequalities. I believe that it is this bridging feature which makes the concept distinctive and particularly useful, both within philosophy and more broadly in society (politics, law, civil society). My hope is that seeing discrimination in this way will allow for a more well-grounded and detailed understanding of our social world, by facilitating a more nuanced application of the concept to specific cases, and by allowing for a richer analysis of their moral and political-philosophical (and potentially even legal) status.

I proceed as follows: in section 2, I propose a definition of discrimination and show how it gives rise to four distinct forms of discrimination. In section 3, I describe cases of implicit bias

discrimination and epistemic injustice, and argue that my definition can capture these on a structural interpretation. This captures the first structural dimension of discrimination. In section 4, I give an account of the moral wrongness of discrimination, in terms of harm. I propose that this account still does not exhaustively capture what makes the phenomenon of discrimination intuitively problematic. I therefore, in section 5, explore the connection between discrimination and theories of social (in)justice. This captures a second, much more foundational, structural dimension of discrimination. Section 6 concludes.

2. The concept of discrimination

So, what is discrimination? In its widest sense, to discriminate is to distinguish or differentiate between things. Obviously, I'm here interested in a much narrower sense. As initially stated, I want to capture the specific phenomenon of discrimination that most of us are concerned with in real life. I propose, tentatively, that this concerns the differential *treatment* of others (paradigmatically: persons), which is in some sense *detrimental* to them, and typically connected to some kind of (perceived) *group* belonging on their part. Moreover, occurrences of the phenomenon typically appear to us as *problematic* (this captures the negative valence contained in utterances such as "But that's discrimination!", or the normative status ascribed to it in legal documents).

I have in previous work proposed a definition of discrimination, close to a number of others' definitions in the literature,[1] but in some respects distinct.[2] This definition aims to accommodate the above concerns:

> *Definition*: An agent, X, (group) discriminates against someone, Y, by φ-ing if and only if:
>> (i) there is a property, P, such that Y has P (or X believes that Y has P),
>> (ii) by φ-ing, X treats Y worse than X would have treated Y, had Y not had P (or had X not believed Y to have P),
>> (iii) it is because Y has P that X treats Y worse by φ-ing (or because X believes that Y has P), and
>> (iv) P is the property of being a member of a socially salient group.

This generic definition can be precisified in different ways, in order to capture several distinct forms of discrimination. To start, consider (ii). This condition can be precisified as:

> (ii') X *would have π-ed*, rather than φ-ed, had Y not (been believed to) have P, and φ-ing toward someone constitutes worse treatment of them than π-ing,

or as:

> (ii") had Y not (been believed to) have P, X *would still have φ-ed*, but φ-ing toward someone with P constitutes worse treatment than φ-ing toward someone without P.

The first way of spelling this out gives us discrimination as *differential treatment*: comparing two different acts, φ-ing vs π-ing. The second gives us discrimination as *disparate impact*:

---

[1] See e.g. (Lippert-Rasmussen 2014) and many of the entries in (Lippert-Rasmussen 2017).
[2] See (Berndt Rasmussen 2019), (Berndt Rasmussen 2020), (Berndt Rasmussen 2023).

comparing the different impacts of one act, φ (under a "facially neutral" description), on someone with P vs without P.

Even condition (iii) can be precisified in two alternative ways, as:

> (iii') It is because *X has P-related intentions* (e.g., X dislikes people with P and believes that Y has P) that X treats Y worse,

or as:

> (iii") It is not because X has P-related intentions that X treats Y worse, but rather because of some *other P-related reason*.

The first way of spelling this out gives us *intentional* discrimination: property P plays a motivational role for X's action. The second gives us *non-intentional* discrimination: property P has an explanatory role with regard to X's action, but not via X's motivational states.

Combining these two pairs of precisifications results in four possible types forms of discrimination. *Table 1* systematises these and exemplifies each form with a paradigmatic case of *race discrimination in enrolment/employment decisions*, where a gatekeeper (X) refuses to accept (φ) an applicant (Y), making the applicant worse off by denying a sought opportunity, just due to the applicant's being black (property P).

	Differential treatment	**Disparate impact**
Intentional	(1) A university in the early 1950's US South accepts a white applicant but turns down an *equally* qualified black applicant, stating: "This is a whites-only university. Blacks are referred to apply to some 'separate-but-equal' university for African Americans."[3]	(2) An employer turns down a *qualified* black applicant, stating: "We don't hire people who lack high school education", while intentionally using this criterion because of its ability to track politically induced, race-correlated educational deficits.[4]
Non-intentional	(3) A university accepts a white candidate for their PhD-programme but turns down an *equally* qualified black candidate, stating the latter was *less* qualified, where the unequitable ranking is due to the evaluators' implicit biases.[5]	(4) An employer turns down a *qualified* black applicant, stating: "We don't hire people who lack high school education", without awareness of the criterion's ability to track politically induced, race-correlated educational deficits.[6]

Table 1: Four forms of discrimination with paradigmatic examples.[7]

---

[3] This example resembles *Sweatt v. Painter*; see (Lavergne 2010). Note that there may but need not be disparate impact under disparate treatment: if (contrary to historical fact) the educational facilities had been separate *and relevantly equal*, blacks might not have been worse off than whites in the labour market, but such non-disadvantageous yet differential treatment would still constitute discrimination and may still be marked as morally wrong as such.

[4] This example resembles *Griggs v. Duke Power Company*; see (Khaitan 2015, 31), but with the addition that the criterion "is covertly used to target members of a protected class" (Mendoza 2017, 258). Cf. even Altman's "Jim Crow era" example (Altman 2016, para. 2.1).

[5] This example may be posited as a specific instantiation of the unequal rankings of identical CVs under different (racially or gender coded) names, which have been extensively studied (Zschirnt and Ruedin 2016). Cf. (Alesina et al. 2018) for the correlation of teachers' implicit anti-immigrant bias and their grading of immigrant vs. native middle school students.

[6] This example resembles *Griggs v. Duke Power Company* under "absence of a discriminatory intent" (Khaitan 2015, 31). There is, of course, the separate but related problem of discrimination at the educational level.

[7] *Table 1* originally in (Berndt Rasmussen 2020, 738).

3. Structural discrimination

Is there conceptual space for structural discrimination in this picture? An answer of course depends on what exactly structural discrimination is. The phenomenon is not very systematically explored in the philosophical literature.[8] One of the more thorough treatments is by Kasper Lippert-Rasmussen (2014). He characterises it broadly as "discrimination without subjects of discrimination, i.e., [without] individual agents who discriminate".[9] This description is then made more precise, as discrimination involving "structural elements" where 'social structure' refers to "a regular patterns of social behavior that supervenes on the behavior or dispositions of the members of the relevant social unit",[10] or to "rules that constitute and regulate the major sectors of life such as family relations, property ownership and exchange, and political powers and responsibilities."[11] Moreover, it "obtains where, and only where, the social structures are such that certain socially salient groups are disadvantaged relative to others and where at least part of the explanation why these structures are in place appeals to the fact that these groups are subjected to or have been subjected to various forms of direct discrimination".[12]

The two main criteria for structural discrimination that I take from these characterisations are:

> (a) there is no specific discriminating agent; rather, the discrimination is a feature of social units;
> (b) its causal mechanism involves social structures: *behavioural patterns* supervening on individual behaviour and dispositions, or *social rules* that constitute and regulate large-scale social interactions, within the relevant social unit.[13]

Criterion (a) appears to be at odds with my definition of discrimination, which relies on there being an agent, X. Of course, such an agent need not be an individual, but can be a collective, e.g., a university and other sorts of institutions, as well as other forms of sufficiently organized collectives.[14] Yet, such forms of discrimination amount to institutional discrimination.[15] Since these retain an agent, they are conceptually distinct from structural discrimination. It could thus be claimed that any phenomenon satisfying criterion (a) cannot be discrimination under my own definition.

---

[8] Structural discrimination as (non-intentional?) disparate impact: "Structural discrimination refers to the policies of majority institutions, and the behavior of the individuals who implement these policies and control these institutions, that are race-neutral in intent but have a differential and/or harmful effect on minority groups." (Pincus 1994)

[9] (Lippert-Rasmussen 2014, 18-19). Note that this is Lippert-Rasmussen's characterization of what others take to be structural discrimination. He holds that "there is *no such thing as agent- or objectless discrimination*" (ibid., italics in the original). However, he allows for a "very thin, grammatical notion of a subject", e.g. a social structure in itself. His own account of discrimination, as well as his further precisifications of structural discrimination (below) are thus compatible with denying that structural discrimination is agent-less in a thicker sense (as in condition (a) and my own definition of structural discrimination, below). Thanks to Søren Flinch Midtgaard for pressing this point.

[10] (Lippert-Rasmussen 2014, 41)

[11] (Lippert-Rasmussen 2014, 78)

[12] Lippert-Rasmussen 2014, ??

[13] These social structures may be (partially) explained by reference to existing or previous direct discrimination of the socially salient groups in question (see above); I will however not rely on or defend this claim here.

[14] Cf. List & Pettit (2011).

[15] Cf. (Lippert-Rasmussen 2014, 77–78)

Together, criteria (a) and (b) also suggest that structural discrimination, while possibly being supervenient on individual intentions that drive individual behaviour, it is necessarily non-intentional, due to lacking an agent. It could thus be claimed that structural discrimination necessarily is "indirect", in the orthodox *non-intentional disparate impact* sense of this term.[16]

In the next two subsections, I challenge both these claims. I explore two mechanisms that give rise to discrimination that is clearly structural, and that fit into my fourfold taxonomy. They are thereby shown to fall under the generic definition. That is to say, the scope of the definition is shown to be broader than its stipulated limitation to 'agents' in the definiendum suggests. Further, some of these mechanisms are shown to give rise to *differential treatment* discrimination, thus falling outside of the orthodox understanding of indirect discrimination.

I make two main arguments: (1) Large-scale behavioural patterns, which make certain socially salient groups worse off within societies, can supervene on *implicit biases* which affect individual dispositions and behaviour without being reducible to specific (individual or collective) agents. (2) The social rules that constitute and regulate large-scale social interactions are in part dependent on shared hermeneutic resources, which are not reducible to specific (individual or collective) agents, and whose distribution may be *epistemically unjust*. These two mechanisms give rise to implicit bias discrimination (section 3.1) and epistemic discrimination (section 3.2), respectively.

3.1 Implicit bias discrimination

Implicit biases can roughly be characterized as stereotypical associations (e.g., "woman–family"; "man–career") or prejudices (e.g., "women aren't good at math") that are held by people against their own explicit convictions. More specifically, they are often understood as some kind of mental states that are automatically activated by social category cues and that are valenced, in the sense of carrying positive or negative evaluations or hierarchical rankings of these social categories. They are moreover typically taken to be beyond their bearer's introspection and direct control – and potentially in conflict with other mental states. And they are taken to be accessible via indirect psychological measures, such as the Implicit Association Test.[17]

Conceptualizing implicit biases as mental states has provoked different sorts of criticism. Empirically, to account for an individual construct, the measures do not seem hold up to scientific standards: individual test results vary over time and test contexts (such that their replicability is in question), and they only weakly correlate with the individual's discriminating behaviour (bringing into doubt the validity of the construct).[18] These problems are exacerbated by findings that individual debiasing interventions only have very short-lived effects,[19] and do not, even when effective, seem to affect discriminating behaviour significantly.[20] Philosophically, attempts to use individualist accounts of implicit bias to

---

[16] Cf. (Lippert-Rasmussen 2014, 78)

[17] These tests are available at https://implicit.harvard.edu/implicit/takeatest.html. The IAT is just one of several measures for assessing implicit biases.

[18] (Oswald et al. 2015).

[19] (Lai et al. 2014); (Lai et al. 2016).

[20] (Forscher et al. 2019). See (Meissner et al. 2019); (Brownstein, Madva, and Gawronski 2019); (Brownstein, Madva, and Gawronski 2020) for empirical literature, but also an insightful critical review of these criticisms

explain social inequalities have been criticized for overlooking or even obscuring social and structural features, thereby obstructing social progress.[21]

Collective accounts of implicit bias can circumvent these criticisms. E.g., the "bias of crowds" account builds on studies that show that aggregated national scores on implicit gender–science bias predict national gender differences in science and mathematics achievements for eighth-graders,[22] and that average regional implicit racial bias among white US residents correlates with racial disparities in the use of lethal police force.[23] On these collective levels, average biases are very stable, and strongly correlated to social outcomes.

The account thus proposes that, rather than reflecting individual mental states, "implicit bias reflects the accessibility of concepts [the likelihood that a thought, evaluation, stereotype, trait, or other piece of information will be retrieved for use] linked to a social category", when such a category is activated.[24] The individual accessibility of such concepts varies with the specific situation of the individual, for example, during the IAT session. Still, when "aggregated across a sample of subjects, the average bias score will reflect the knowledge [of stereotypes and prejudices] with the most widely shared accessibility", while idiosyncratic variations "will be averaged away because they are randomly distributed across persons".[25] Thus, individual implicit bias should be understood as "a psychological marker of systemic prejudice in the environment".[26]

The "bias of crowds" model takes into account the empirical criticism against individualist accounts of implicit bias, and instead analyses the phenomenon on a collective level where the measurements are both stable over time and strongly correlated to relevant outcomes (social inequalities). Moreover, the model opens new venues for responding to the philosophical criticism, by moving our attention from bias-reduction by individual intervention (exploited by commercial bias training programs that merely serve as pr-strategies) toward structural solutions (changing our social world by more radical measures, such as affirmative action and institutional redesign).

However, if implicit bias is "a psychological marker of systemic prejudice in the environment", a question arises: who, exactly, are the agents of discrimination and where, exactly, does the discrimination take place? How, more generally, should we understand the mechanisms linking implicit biases to the social inequalities, e.g., in the labour market, that we are concerned with? In a recent article, my co-author and I propose such a general mechanism, by combining Cristina Bicchieri's game theoretic framework of social norms with empirical findings from implicit bias research. Our model helps explain the puzzling coexistence of racial inequalities in real-world job markets and persisting egalitarian (non-racist) norms. The model consists of a dynamic "hiring game" of explicitly egalitarian-minded norm followers, who nevertheless have implicit biases, along the following lines.[27]

Imagine a big firm consisting of a large number of subsections, each headed by a manager. From time to time these managers hire new people. Assume that the firm is initially all-

---

[21] (Haslanger 2015).
[22] (Nosek et al. 2009, 10596).
[23] (Hehman, Flake, and Calanchini 2018). For further studies, see (Payne, Vuletich, and Lundberg 2017).
[24] (Payne, Vuletich, and Lundberg 2017, 235).
[25] (Payne, Vuletich, and Lundberg 2017, 237)
[26] (Payne, Vuletich, and Lundberg 2017, 238)
[27] (Berndt Rasmussen and Olsson Yaouzis 2023).

white. Managers know that there is competence among the black applicants which would benefit the firm overall. Still, each prefers to head a racially homogenous subsection, because it saves them the trouble of (supposed) interracial conflicts in their own workplace. They are thus all trapped in a prisoner's dilemma: each doing what is (supposedly) better for them (their own subsection), while they all (the firm) miss out on competence.

On Bicchieri's game-theoretic account, social norms solve such prisoner's dilemmas.[28] To illustrate: suppose that the managers become aware of an egalitarian social norm: "When hiring, hire the most competent candidate, regardless of their race". This norm changes their incentives: as long as they believe that enough others will both comply with it and expect *them* in turn to comply, they want to comply with it themselves. Imagine now that each manager comes to believe this about the other managers. Each then complies with the egalitarian norm and hires black applicants whenever they are the most competent. Slowly, the racial composition of the firm will change.

However, if there are implicit racial biases among the managers, this will sometimes distort their actions. They want to comply with the norm but sometimes make mistakes. Due to the nature of implicit bias these mistakes are asymmetrical. That is, they sometimes occur when the most qualified candidate is black (such that a white candidate is hired instead), but never occur when the most qualified candidate is white (such that a black candidate would be hired).

Now, assume that there are many decisions and many managers, so mistakes add up. This could, on the whole, explain large scale hiring inequalities. But this would mean that each of them, observing that the egalitarian norm is violated time and again, would cease to believe that enough others complied with it. Moreover, observing that such frequent norm violations are not met with protests by the others, each manager would cease to believe that enough others expect *them* in turn to comply. Then, each would no longer want to comply and the norm would break down.

Yet this is not what seems to happen in the real world. Rather, the norm stays in place (people uphold the egalitarian norm when surveyed) *and* large-scale hiring discrimination persists (causing pervasive racial inequalities). Our model can account for this, by illuminating the intricate interplay between implicit bias and job competence. To see this, consider a specific recruitment case, where the most competent candidate is black. Their race is a clearly observable feature – their competence typically is not.

Suppose that the hiring manager makes an implicit bias-mistake and hires a less competent white candidate. The other managers likely cannot directly observe that a norm violation has taken place. They can, however, observe the successful candidate's race. If they (like most of us) hold implicit racial biases, they may perceive the white candidate as more competent than they actually are, and (falsely) infer that the egalitarian norm was followed. Thus no one protests, and no one changes their belief that enough others comply with the norm and expect them in turn to comply. The norm may be repeatedly violated, but does not break down. In such a scenario, there may be severe racial inequalities, even though all agents are egalitarian norm-followers, in the sense that they are motivated by egalitarian (non-racist or anti-racist) beliefs and desires.[29]

---

[28] (Bicchieri 2005).

[29] There is a further interesting feature of implicit bias that is brought out by our model: suppose the manager *doesn't* make a mistake. The most competent black candidate is hired. Again, the others can observe the

If this picture is correct, the systematic non-hiring of black applicants amounts to *differential treatment* discrimination, since we here compare the act of hiring/enrolling of applicants with the act of rejecting them, where the choice (or rather, its probabilities) depends on a socially salient feature, i.e., the applicants' race. Moreover, the non-hiring is *non-intentional*, since race does not enter the relevant motivational (belief–desire) states of the individual gatekeepers. In fact, the systematic non-hiring emerges not on the level of individual decisions but on the collective (statistical) level of differential hiring chances. Yet on this level, there is no collective agent (with sufficiently coordinated joint individual intentions to act as a group), but just a collection of individuals whose individual probabilities to make implicit bias-mistakes "happen to" form a systematic pattern.[30]

In light of this, example (3) in *Table 1* should thus be redescribed as follows:

> (3') In the context of university admissions, white candidates for PhD-programmes are significantly more likely to be accepted than equally qualified black candidates, where the unequitable chances are due to (and correlate with) aggregate measures of implicit racial biases in the relevant social context.

This is a clear example of structural discrimination, in the sense that (a) there is no specific discriminating agent; rather, the discrimination is a feature of social units; and (b) its causal mechanism involves social structures: *behavioural patterns* supervening on individual behaviour and dispositions. At the same time, (3') is captured by the above definition of discrimination, as non-intentional differential treatment, falling under conditions (ii') and (iii'').

3.2 Epistemic discrimination

Being a knower is an integral part of being human. Social practices of knowledge acquisition and transfer are all around us, and taking part in them as knowers matters to people. According to Miranda Fricker's (2007) pathbreaking account, epistemic injustice is a "distinctively epistemic kind of injustice", viz., "a wrong done to someone specifically in their capacity as a knower", which occurs because of "identity prejudice", i.e., "prejudice relating to social identity".[31] In this section, I analyse several forms of such wrongings as forms of discrimination, with a special eye to when they can be said to constitute structural discrimination. I consider three examples: Fricker's own two forms of epistemic injustice, testimonial and hermeneutical injustice, and the phenomenon of stereotype threat (since it offers an interesting addition to these standard examples, in the context of structural discrimination).

---

candidate's race but not their competence. Again, if they hold implicit racial biases, they may perceive the black candidate as less competent than they actually are – and (falsely) infer that the egalitarian norm was violated. If this happens repeatedly, observers might eventually (falsely) conclude that the egalitarian norm has come to be replaced with a norm of "political correctness": "When hiring, hire the most qualified minority candidate (to increase firm diversity)". The model thus helps us see that implicit bias plays a double role in hiring contexts: it affects the hiring decisions themselves, but also the bystander evaluations of these decisions. It thereby illuminates how hiring discrimination can be invisible in seemingly egalitarian social contexts — and how instead *non-discrimination* may appear suspicious.

[30] Cf. Of course, in some cases there might be such collective agents whose decisions are affected by the aggregated implicit biases of its members. In such cases, the discrimination is not structural, but institutional (see above).

[31] Fricker (2007).

Testimonial injustice occurs when a speaker suffers a credibility deficit because of an identity prejudice. That is, in a situation of knowledge transfer through testimony, they are ascribed less credibility than they are due, due to their group membership. An obvious example is a sexual assault trial where the female victim's looks and manners are utilised to discredit her own words. Such credibility deficits can arise from the denial that the speaker is knowledgeable (a capable knower), or trustworthy (a sincere knowledge transmitter).

Such cases are, then, always instances of *differential treatment*, since they involve, on part of the hearer, an act of assigning reduced credibility vs an act of assigning due credibility.[32] They may be instances of *intentional* differential treatment, and thus be classified as non-structural discrimination, if the identity prejudice in question is explicit. Some cases may, on the other hand, involve identity prejudice in the form of implicit bias. There is evidence that, e.g., for student course evaluations, male professors are rated as more knowledgeable than their female counterparts,[33] and judgments of trustworthiness have been shown to correlate with implicit race attitudes.[34] Implicit biases may thus affect the degree of credibility ascribed to marginalised groups. Such cases of testimonial injustice may then be classified as *testimonial implicit bias discrimination*, and thus constitute structural discrimination, along the above lines of reasoning.

The aptness of an analysis on a structural level becomes even more clear when we consider a specific form of testimonial injustice, dubbed as *testimonial smothering*. This is a form of self-silencing, which occurs when a speaker perceives their "immediate audience as unwilling or unable to gain the appropriate uptake of proffered testimony".[35] Suppose this unwillingness or inability of the audience has to do with some property P, viz., the speaker's socially salient group membership. On a non-structural analysis, such a case would risk being classified as self-discrimination: the speaker's choosing to refrain from stepping forth in her capacity as a knower (i.e., choosing to withhold, rather than share, her testimony). This would then arguably not be due to the speaker's own P-related intentions (of, e.g. disliking or devaluing her own social group), but rather because of another P-related reason: that the audience lacks the opportunity for an appropriate uptake of the testimony, due to the speaker's social group membership. At first sight, it seems we would deal with a case of non-intentional differential (self-)treatment, falling under conditions (ii') and (iii'') on a non-structural, agent-based reading.

This analysis is obviously flawed, not least in singling out the *speaker* as a discriminator. However, a structural reading of this case allows us to locate the problem in its proper place: in the audience's inability to gain the appropriate uptake, which is due to P-related (explicit or implicit) prejudices. The discrimination is structural, in the sense of being a feature of the social unit in question: the behavioural patterns supervening on individual behaviour and dispositions (or the social rules that constitute and regulate large-scale social interactions) in the surrounding society, are not conducive to the appropriate uptake. This feature, known by the would-be speaker, makes her unable to testify altogether. If this lack of uptake on part of the audience is due to the speaker's social group membership (such that, had she not had P there would have been uptake), we are dealing with *non-intentional differential treatment* of a structural kind.

---

[32] I here (for now) bracket cases of credibility excess.
[33] (Boring 2017).
[34] (Charbonneau et al. 2020).
[35] (Dotson 2011, 244).

Hermeneutical injustice occurs when there is a lack of shared concepts for understanding and communicating the social experiences of marginalised groups. The standard example concerns the concept of sexual harassment. Although the phenomenon itself runs through all of recorded human history, the term was coined in separatist women's groups in the 1960s. Filling a hermeneutical gap, this concept then allowed for a new understanding of the shared patterns, immense scope, and indeed long history of myriads of individual "nuisances" women experienced (and continue to experience) in varying contexts.

When shared hermeneutical resources are missing, both dominant and marginalised groups are equally affected in one sense: the hermeneutical gap is a cognitive obstacle to articulating and analysing certain social phenomena for anyone alike. However, the gap is prudentially and socially disadvantageous for marginalised groups (while neutral or even advantageous for the dominant group), by making it harder to communicate their specific social experiences and to coordinate with others to address them. This characterisation supports an analysis of instances of hermeneutical injustice as *disparate impact* discrimination: people lack the concept of sexual harassment irrespective of their social group membership, but lacking this concept as a woman (being far more likely to be subjected to acts captured by it) makes one worse off than lacking this concept as a man. However, the hermeneutical gap is not any kind of "treatment" that an agent (individual or collective) directs against another. Rather, the causal mechanism runs via the shared social – linguistic – rules that constitute and regulate large-scale social interactions. There is no specific agent, rather, the discrimination in question a feature of the relevant social unit. This means that we are clearly dealing with structural discrimination, according to (a) and (b) above. It also means that this phenomenon can be captured by the above definition of discrimination, as non-intentional disparate impact, falling under conditions (ii'') and (iii'').

(Contrast such standard cases of hermeneutical discrimination to the related phenomena of *contributory injustice*[36] and *wilful hermeneutical ignorance*.[37] Here, the conceptual resources exist but are ignored by dominant social groups, due to P-related intentions. When such wilful ignorance makes an audience fail in their uptake of testimony, we are dealing with intentional disparate impact, i.e., agent-based, non-structural discrimination, falling under conditions (ii'') and (iii').)

Stereotype threat occurs when one's epistemic performance is thwarted due to the activation of stereotypes about one's own group.[38] E.g., when gender stereotypes such as "women are bad at math" are activated prior to a math test, women tend to underperform. This may be due to increased anxiety – about one's own mathematical aptitude and about confirming the stereotype, should one do poorly on the test. Stereotype activation can be overt (stated out loud) or covert (one simply finds oneself as the only woman in the room) and may target one's own explicit prejudices or implicit biases (such as measured by the Science–Gender IAT).[39]

On a non-structural analysis – just as in the above case of testimonial smothering – such a case of stereotype threat risks being classified as self-discrimination. In this case we would have to say that the agent, by performing worse on the test than her abilities allow, acts in a

---

[36] Dotson 2012.
[37] Polhaus 2012.
[38] (Spencer, Logel, and Davies 2016).
[39] (Greene 2020).

way that makes things worse for herself. Had she not been a woman, the stereotype would not have made her anxious, and she would have performed better. Again, this would arguably not be intentional, but rather due to P-related reasons (anxiety activated by P-related stereotypes). Still, at first sight, it seems we would again deal with a case of self-discrimination, in the form of non-intentional differential treatment (in a non-structural sense).

Again, this analysis is obviously flawed. The discrimination should be analysed as structural, in the sense of being a feature of the social unit in question: the society which surrounds the agent, and in which such gender stereotypes are widely spread (leading to possible activation in, e.g., a math test situation) and internalised by the agent (leading to activation of anxiety). We are then dealing with non-intentional differential treatment, falling under conditions (ii') and (iii'') on a structural reading. *Table 2* maps the different forms of epistemic injustice onto the conceptual framework of discrimination and provides a paradigmatic example for each.

	Differential treatment	Disparate impact
Intentional (agential)	(1) In a sexual assault trial, the perpetrator's lawyer utilises the female victim's looks and manners to discredit her own words (where the identity prejudice in question is explicit): *agential testimonial injustice.*	(2) When reporting a sexual harassment incident to her boss, a female employee is met with denial and laughter: "He was just flirting with you": *contributory injustice / wilful hermeneutical ignorance.*
Non-intentional *in a structural sense*	(3) In jury trials, black defendants are significantly more likely to be assigned credibility deficits than white defendants, where the inequality is due to (and conviction rates correlate with) aggregate measures of implicit racial biases in the relevant social context: *structural testimonial injustice.* (4) In an academic environment that communicates gender stereotypes (e.g. via the wall of fame picturing all the "great philosophers"), female students tend to underperform: *stereotype threat.*	(5) In a culture that lacks the concept of sexual harassment, women *and* men suffer a cognitive disadvantage, but women are made worse off than men, since they are far more likely to be subjected to the phenomenon: *structural hermeneutical injustice.* (6) In a work environment where the officials are unwilling or unable to gain the appropriate uptake for testimonies of sexual harassment, women (who are more likely to be subjected to it) are less likely to call it out and get redress: *testimonial smothering.*

Table 2: Forms of epistemic discrimination with paradigmatic examples.

3.3. Upshots

To sum up, sections (3.1) and (3.2) analyse cases of structural discrimination, as characterised by criteria (a) and (b) above. They are shown to fall under the above generic definition of discrimination, in the sense that they satisfy some combinations of its precisified conditions (ii) and (iii) (as well as, it has been assumed, the other conditions). This means that the scope of the definition of discrimination is broader than its stipulated limitation to 'agents' in the definiendum suggests. That is, the conditions stated in the definiens may be satisfied in cases where there is no agent, in an individual or collective sense. The limitation to 'an *agent*, *X*' in

the definiendum is thus arbitrary and can – without loss of analytical precision, but with potential gains from broader scope – be changed to 'an *entity*, X':

> *Definition**: An entity, X, (group) discriminates against someone, Y, by φ-ing if and only if:
>> (i) there is a property, P, such that Y has P
>> (or is perceived to have P),
>> (ii) by φ-ing, X treats Y worse than X would have treated Y, had Y
>> not had P (or had not been perceived to have P),
>> (iii) it is because Y has P that X treats Y worse by φ-ing
>> (or because Y is perceived to have P), and
>> (iv) P is the property of being a member of a socially salient group.

I will, in the rest of this paper, rely on this improved definition. One structural dimension of discrimination is thus that there can be structural *forms* of differential treatment or disparate impact, that are classified as discrimination by the generic definition.

A further upshot is that the claim that all structural discrimination is non-intentional, and therefore "indirect" (in a conventional sense),[40] needs to be handled with care. If by 'indirect discrimination' we just mean 'non-intentional differential treatment *or* disparate impact', this claim is supported by my analysis. But if by 'indirect discrimination' we mean 'disparate impact' (of a "facially neutral policy",[41] as it is often framed), the claim is not supported. Rather, my analysis shows that structural discrimination can consist of differential treatment (e.g., *implicit bias discrimination* or *stereotype threat*) or disparate impact (e.g., *structural hermeneutical injustice* or *testimonial smothering*).

Finally, analysing many cases of epistemic injustice, as well as implicit bias discrimination, as cases of structural discrimination speaks in favour of addressing the problem (at the very least in these cases) as a social or structural phenomenon, rather than one of individual wrongdoing (or, indeed, individual epistemic virtue or vice, pace Fricker[42]). What is needed is structural change: to courtroom, workplace, and academic cultures; and to society at large in order to change implicit biases (e.g., through affirmative action or other more radical measures).

4. The moral wrongness of discrimination

The definition of discrimination, as it stands, is non-moralised. But it brings out a normative feature which should be captured by any plausible account of the moral wrongness of discrimination: that Y is subjected to worse treatment by X, in some sense, as stated by condition (ii). This allows for different accounts of moral wrongness to be plugged in here, spelling out worse treatment in terms of, e.g., disrespect,[43] demeaning,[44] freedom,[45] or harm.[46]

---

[40] (Lippert-Rasmussen 2014, 78); see (Berndt Rasmussen 2020) for a critique of the conventional distinction.
[41] Cf. Altman (2016, para. 3.1); Moreau (2017, pp. 166–167); …
[42] (Fricker 2007); cf. (Anderson 2012).
[43] Eidelson 2015, cf. Beeghly 2017.
[44] Hellman 2008, 2017; Scanlon 2008.
[45] Moreau 2010, 2012.
[46] Arneson 2006; Lippert-Rasmussen 2014.

I have, in a previous paper,[47] argued for a counterfactual-harm based account – albeit one that appeals to an unorthodox, broad concept of harm. This concept has a welfarist component, of being made worse off in terms of welfare (capturing the orthodox sense of harm[48]), but also a non-welfarist component, of being treated as inferior. The latter component captures features in the vicinity of the above-mentioned disrespect and demeaning accounts; thus, my own approach could be spelled out as a hybrid harm-*and-inferior*-treatment (in the relevant sense) approach. The issue is mainly terminological, although I believe that there is something to be said for the idea that being treated as inferior can reasonably be seen as a kind of (relational) harm to the individual. I will, in any case, not defend this harm-based account here, but merely stipulate it as a basis for further discussions of the structural dimensions of discrimination:

> *Moral wrongness*: An instance of (group) discrimination, φ, by X against Y, on grounds of P, is (prima facie) wrong because X by φ-ing treats Y worse, in the following sense:
> (a) making Y worse off, or
> (b) treating Y as inferior,
> than X would have, had Y not had P (or had Y not been perceived to have P).

To illustrate the advantage of employing an unorthodox, broad harm-account (or a hybrid account), consider two cases of epistemic injustice that are instances of prejudice-based credibility excess. These are cases of knowledge transfer through testimony, where the speaker is ascribed *more* credibility than they are due, because of their group membership. At first sight, this may seem to be advantageous to the speaker. Yet it can turn out to be a clear-cut cut case of wrongful discrimination:

> If I take a black man to be highly knowledgeable about, say, guns or drugs simply because he is a black man, then he has been wronged as a knower just as much as if I take him to be completely ignorant of Shakespeare. Being regarded as highly knowledgeable about domains that are stigmatized or devalued can be insulting to a speaker qua knower, regardless of any long-term disadvantages that might be accrued. [… He] rightly finds it disrespectful with respect to his epistemic identity.[49]

Here, the non-welfarist part of the wrongness account can cover the moral fault of a prejudice-based *credibility excess in stigmatized domains* – even if the speaker is not made worse off in welfarist terms.

Consider a second case of prejudice-based credibility excess:

> [In the US legal system] confessions are often acquired through coercion, manipulation, and deception – such as by using maximization tactics and presenting false evidence – as well as by targeting vulnerable suspects – such as those who have been sleep deprived, juveniles, and the mentally impaired. […]

---

[47] (Berndt Rasmussen 2019)
[48] Cf. (Parfit 1986, 487).
[49] (Lackey 2018).

> when the testimony of a confessing self is privileged over a recanting self because of prejudice, whether racial or otherwise, this results in a unique kind of testimonial injustice that is due to a credibility excess.[50]

Here, the welfarist part of the harm account clearly covers the moral fault of prejudice-based credibility excess that leads to false conviction (irrespective of whether the assignment of the credibility excess itself also involves inferior treatment). The devastating outcome of being falsely sentenced clearly makes the suspect worse off.

Note that the suggested wrong-making feature can capture the wrongness of structural forms of discrimination. It focuses on the discriminatee, Y, who might be an individual or a group (supposing that we can make sense of groups being made worse off or treated as inferior). It can thus capture cases where there is no individual or collective agent, X, who discriminates, but rather a social entity involving social structures (which is something that, e.g., mental state accounts of the wrongness of discrimination would struggle with).

Note, further, that the definition of discrimination, while it is formulated in evaluative terms of 'treating worse', does not make any moral claim. Such a claim is introduced by the separate and non-definitional account of the wrong-making feature of discrimination. The definition is thus non-moralised. The theory in its entirety of course is not – and should not be. Recall that I set out to capture the specific phenomenon of discrimination that most of us are concerned with in real life: as a phenomenon that typically appears to us as *problematic*. Now, my definition carves out a very specific social phenomenon (group discrimination, in an agential or structural sense), which most of us take issue with. The wrongness account then merely spells out why we are morally justified in taking issue with it, i.e., why it is an apt candidate for at least *prima facie* moral wrongness.

However, I now want to bring out that the moral wrongness account does not exhaust the problematic features of discrimination. Yes, discrimination is (prima facie) wrong because it constitutes a harm to the discriminatee, but the problem does not stay there. The harm is not just any kind of arbitrary harm; it is harm due to the discriminatee's (perceived) socially salient feature. This is what makes discrimination especially pernicious, and not just in a moral, but also a social-political sense. This should be brought out and analysed by a theory of discrimination. I briefly sketch this idea in the next section, in order to relate it to the issue of structural discrimination.

5. Discrimination and injustice

Until now, I have not considered conditions (i) and (iv) of the definition of discrimination:

> (i) there is a property, P, such that Y has P (or is perceived to have P), and
> (iv) P is the property of being a member of a socially salient group.

To know where and when these conditions apply, we need to know what 'socially salient group' means. A standard way to understand the expression is this:

---

[50] (Lackey 2021, para 3).

> A group is socially salient if [and only if] perceived membership of it is important to the structure of social interactions across a wide range of social contexts.[51]

This seems to be on the right track, but we can dig deeper at this point. It is not any structures of social interactions that are relevant, but rather the ones that are arise from systematic, unjustified social inequalities – those which theories of distributive social justice and injustice help us identify.[52] What makes discrimination especially problematic is that the harm done to discriminatees latches onto, reproduces, and over time exacerbates such social inequalities. Specifically, we take issue with instances of such harms when they are done to individuals (or groups) who are already on the systematically disadvantaged side of the social inequalities in question – and where this is, moreover, *due to* them belonging to this side in the first place. To capture this, I propose the following analysis of 'socially salient group':

> A group is socially salient if and only if its members are subject to systematic, unjustified social disadvantages within the given social context.[53]

This analysis allows us to spell out a deeper problematic feature of discrimination. It is not only morally wrong but also unjust:

> *Injustice*: Discrimination is unjust because it reproduces and exacerbates systematic, unjustified social disadvantages.

Given this analysis, we can now plug in our preferred theory of distributive social justice into the theory of discrimination (just as we did with our preferred account of moral wrongness – here: the broad harm account). Such a theory is needed to spell out the relevant distributive patterns and the relevant currency of justice. It can then serve to provide us with the criteria to find whatever are the specific *grounds* of discrimination in a given context. In many of today's societies, these grounds will be the well-known categories of race, gender, disability, religion, and so on. And such exact categories arguably need to be specified, e.g., in legal statutes and institutional policies, in order to allow for the efficient and orderly application of such rules in real-life situations. However, behind any such specific categories, there should be a general criterium.[54] This should be brought out by a theory of discrimination, not least so that we, from time to time, can re-evaluate the specific grounds that happen to be codified in our legal statutes and institutional policies.

---

[51] (Lippert-Rasmussen 2014, 30).

[52] E.g., (Rawls 1999).

[53] I here want to note (although I do not have space to develop and defend) the upshot that this limits discrimination to socially disadvantaged groups. Those on the systematically *advantaged* side of the social inequalities in question can thus not be discriminated against, according to my account. This may seem counterintuitive at first, but arguably has the advantage of classifying, e.g., affirmative action (of the kind just recently ruled out as unconstitutional by the US Supreme Court) as non-discrimination. (Cf. https://www.washingtonpost.com/politics/2023/06/29/affirmative-action-supreme-court-ruling/).

[54] Such a criterion need of course not solely pick out single-dimensional categories. This approach may thus provide a new pathway for taking on the intersectionality challenge (a possibility that I'm hoping to explore further).

5.1 Upshots

A normative upshot is that this framework brings out that discrimination is, in fact, doubly problematic: it is a driver of social injustice, which operates by wronging (harming) already disadvantaged individuals – adding, as it were, injustice to injury.

A further noteworthy upshot concerns the debate on epistemic injustice. The present framework of discrimination firmly anchors problems discussed within social epistemology (e.g., testimonial and hermeneutical injustice, which are in themselves clearly epistemic in nature) within a social justice framework. It thereby shows us when, why, and how, more exactly, epistemic injustice matters so much in social contexts.

A conceptual upshot is that we should see discrimination as a bridge concept. As such, it is distinctive and particularly useful exactly because it connects wrongness to injustice: the moral domain (where individuals can be harmed, and thus *prima facie* wronged) to the political domain (where there are unjustified inequalities between social groups, i.e., social injustice).

Moreover, providing a theory of discrimination that connects its moral wrongness to social injustice, the framework brings out a second, and indeed foundational, structural dimension of discrimination. The analysis of 'social salience' captures a structural social phenomenon par excellence: a feature of social units consisting of distributive patterns of "primary goods", resulting from (without being reducible to) individual behaviour and dispositions. Notably, this structural dimension is derived from an analysis of 'social salience', which is part of the conditions of a *generic* definition of discrimination. In this sense, *all* forms of discrimination, be they structural or agential, are a structural phenomenon.

6. Conclusions

In this paper, I have explored ways to capture the structural dimensions of discrimination. I have outlined a theory of discrimination that aims to capture what many of us are concerned with in real life cases, by providing a definition as well as an account of its wrong- and unjust-making features. I have proposed a taxonomy of four distinct forms of discrimination and explored how implicit bias discrimination and epistemic discrimination fit into this frame, arguing that my definition can capture these as discrimination if we allow for a structural reading. This captured a first structural dimension of discrimination. I have moreover proposed that an account of the moral wrongness of discrimination, in terms of a broad concept of harm, does not exhaustively capture what makes the phenomenon of discrimination intuitively problematic. I have therefore further explored the connection between discrimination and theories of social (in)justice. This has, finally, revealed a second, much more foundational, structural dimension of discrimination – one that applies to all forms of discrimination, be they structural or agential.[55]

References

Alesina, Alberto, Michela Carlana, Eliana La Ferrara, and Paolo Pinotti. 2018. 'Revealing Stereotypes: Evidence

---

[55] Thanks to participants of the 2024 Swedish Congress of Philosophy (University of Gothenburg, June 2024) and of the CEPDISC workshop on "Structural Discrimination" (Aarhus University, October 2023), and especially to Tom Parr, for valuable comments on earlier drafts of this paper.

from Immigrants in Schools'. Working Paper 25333. National Bureau of Economic Research. https://doi.org/10.3386/w25333.

Altman, Andrew. 2016. 'Discrimination'. In *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta, Winter 2016. Metaphysics Research Lab, Stanford University. https://plato.stanford.edu/archives/win2016/entries/discrimination/.

Anderson, Elizabeth. 2012. 'Epistemic Justice as a Virtue of Social Institutions'. *Social Epistemology* 26 (2): 163–73. https://doi.org/10.1080/02691728.2011.652211.

Berndt Rasmussen, Katharina. 2019. 'Harm and Discrimination'. *Ethical Theory and Moral Practice* 22 (4): 873–91. https://doi.org/10.1007/s10677-018-9908-4.

———. 2020. 'Implicit Bias and Discrimination'. *Theoria* 86 (6): 727–48. https://doi.org/10.1111/theo.12227.

———. 2023. 'Age Discrimination: Is It Special? Is It Wrong?' In *Ageing without Ageism?: Conceptual Puzzles and Policy Proposals*, edited by Greg Bognar and Axel Gosseries, 0. Oxford University Press. https://doi.org/10.1093/oso/9780192894090.003.0002.

Berndt Rasmussen, Katharina, and Nicolas Olsson Yaouzis. 2023. 'The Tyranny of Political Correctness? A Game-Theoretic Model of Social Norms and Implicit Bias'. *Journal of Applied Philosophy* n/a (n/a). https://doi.org/10.1111/japp.12690.

Bicchieri, Cristina. 2005. *The Grammar of Society: The Nature and Dynamics of Social Norms*. Cambridge University Press.

Boring, Anne. 2017. 'Gender Biases in Student Evaluations of Teaching'. *Journal of Public Economics* 145 (January):27–41. https://doi.org/10.1016/j.jpubeco.2016.11.006.

Brownstein, Michael, Alex Madva, and Bertram Gawronski. 2019. 'What Do Implicit Measures Measure?' *WIREs Cognitive Science* 10 (5): e1501. https://doi.org/10.1002/wcs.1501.

———. 2020. 'Understanding Implicit Bias: Putting the Criticism into Perspective'. *Pacific Philosophical Quarterly* Early View. https://doi.org/10.1111/papq.12302.

Charbonneau, Isabelle, Karolann Robinson, Caroline Blais, and Daniel Fiset. 2020. 'Implicit Race Attitudes Modulate Visual Information Extraction for Trustworthiness Judgments'. *PLOS ONE* 15 (9): e0239305. https://doi.org/10.1371/journal.pone.0239305.

Dotson, Kristie. 2011. 'Tracking Epistemic Violence, Tracking Practices of Silencing'. *Hypatia* 26 (2): 236–57. https://www.jstor.org/stable/23016544.

Forscher, Patrick S., Calvin K. Lai, Jordan R. Axt, Charles R. Ebersole, Michelle Herman, Patricia G. Devine, and Brian A. Nosek. 2019. 'A Meta-Analysis of Procedures to Change Implicit Measures'. *Journal of Personality and Social Psychology* 117 (3): 522–59. https://doi.org/10.1037/pspa0000160.

Fricker, Miranda. 2007. *Epistemic Injustice: Power and the Ethics of Knowing*. Oxford University Press. http://www.oxfordscholarship.com/view/10.1093/acprof:oso/9780198237907.001.0001/acprof-9780198237907.

Greene, Nathifa. 2020. 'Stereotype Threat, Identity, and the Disruption of Habit'. In *An Introduction to Implicit Bias*. Routledge.

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

Hehman, Eric, Jessica K. Flake, and Jimmy Calanchini. 2018. 'Disproportionate Use of Lethal Force in Policing Is Associated With Regional Racial Biases of Residents'. *Social Psychological and Personality Science* 9 (4): 393–401. https://doi.org/10.1177/1948550617711229.

Khaitan, Tarunabh. 2015. *A Theory of Discrimination Law*. Oxford University Press. http://www.oxfordscholarship.com/view/10.1093/acprof:oso/9780199656967.001.0001/acprof-9780199656967.

Lackey, Jennifer. 2018. 'Credibility and the Distribution of Epistemic Goods'. In *Believing in Accordance with the Evidence: New Essays on Evidentialism*, edited by Kevin McCain, 145–68. Synthese Library. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-95993-1_10.

———. 2021. 'False Confessions and Subverted Agency'. *Royal Institute of Philosophy Supplements* 89 (May):11–35. https://doi.org/10.1017/S1358246121000072.

Lai, Calvin K., and et al. 2014. 'Reducing Implicit Racial Preferences: I. A Comparative Investigation of 17 Interventions.' *Journal of Experimental Psychology: General* 143 (4): 1765–85. https://doi.org/10.1037/a0036260.

Lai, Calvin K., Allison L. Skinner, Erin Cooley, Sohad Murrar, Markus Brauer, Thierry Devos, Jimmy Calanchini, et al. 2016. 'Reducing Implicit Racial Preferences: II. Intervention Effectiveness across Time'. *Journal of Experimental Psychology: General* 145 (8): 1001–16. https://doi.org/10.1037/xge0000179.

Lavergne, Gary M. 2010. *Before Brown: Heman Marion Sweatt, Thurgood Marshall, and the Long Road to Justice*. University of Texas Press.

Lippert-Rasmussen, Kasper. 2014. *Born Free and Equal?: A Philosophical Inquiry into the Nature of Discrimination*. Oxford; New York: Oxford University Press.

———. 2017. *The Routledge Handbook of the Ethics of Discrimination*. Routledge Handbooks in Applied Ethics. New York: Routledge. https://ezp.sub.su.se/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=nlebk&AN=1581310&site=ehost-live&scope=site.

Meissner, Franziska, Laura Anne Grigutsch, Nicolas Koranyi, Florian Müller, and Klaus Rothermund. 2019. 'Predicting Behavior With Implicit Measures: Disillusioning Findings, Reasonable Explanations, and Sophisticated Solutions'. *Frontiers in Psychology* 10 (November). https://doi.org/10.3389/fpsyg.2019.02483.

Mendoza, José Jorge. 2017. 'Discrimination and Immigration'. In *The Routledge Handbook of the Ethics of Discrimination*, edited by Kasper Lippert-Rasmussen. London ; New York: Routledge.

Nosek, Brian A., Frederick L. Smyth, N. Sriram, Nicole M. Lindner, Thierry Devos, Alfonso Ayala, Yoav Bar-Anan, et al. 2009. 'National Differences in Gender–Science Stereotypes Predict National Sex Differences in Science and Math Achievement'. *Proceedings of the National Academy of Sciences* 106 (26): 10593–97. https://doi.org/10.1073/pnas.0809921106.

Oswald, Frederick L., Gregory Mitchell, Hart Blanton, James Jaccard, and Philip E. Tetlock. 2015. 'Using the IAT to Predict Ethnic and Racial Discrimination: Small Effect Sizes of Unknown Societal Significance'. *Journal of Personality and Social Psychology* 108 (4): 562–71. https://doi.org/10.1037/pspa0000023.

Parfit, Derek. 1986. *Reasons and Persons*. Oxford, New York: Oxford University Press.

Payne, B. Keith, Heidi A. Vuletich, and Kristjen B. Lundberg. 2017. 'The Bias of Crowds: How Implicit Bias Bridges Personal and Systemic Prejudice'. *Psychological Inquiry* 28 (4): 233–48. https://doi.org/10.1080/1047840X.2017.1335568.

Pincus, Fred L. 1994. 'From Individual to Structural Discrimination'. In *Race And Ethnic Conflict*. Routledge.

Rawls, John. 1999. *A Theory of Justice: Revised Edition*. Cambridge, MA: Belknap Press.

Spencer, Steven J., Christine Logel, and Paul G. Davies. 2016. 'Stereotype Threat | Annual Review of Psychology'. *Annual Review of Psychology* 67:415–37. https://doi.org/10.1146/annurev-psych-073115-103235.

Zschirnt, Eva, and Didier Ruedin. 2016. 'Ethnic Discrimination in Hiring Decisions: A Meta-Analysis of Correspondence Tests 1990–2015'. *Journal of Ethnic and Migration Studies* 42 (7): 1115–34. https://doi.org/10.1080/1369183X.2015.1133279.