

# Harm and Discrimination

Katharina Berndt Rasmussen<sup>1,2</sup> 

Accepted: 5 July 2018 / Published online: 1 August 2018  
© The Author(s) 2018

**Abstract** Many legal, social, and medical theorists and practitioners, as well as lay people, seem to be concerned with the harmfulness of discriminative practices. However, the philosophical literature on the moral wrongness of discrimination, with a few exceptions, does not focus on harm. In this paper, I examine, and improve, a recent account of wrongful discrimination, which divides into (1) a definition of group discrimination, and (2) a characterisation of its moral wrong-making feature in terms of harm. The resulting account analyses the wrongness of discrimination in terms of intrapersonal comparisons of the discriminatee’s actual, and relevantly counterfactual, well-being levels. I show that the account faces problems from counterfactuals, which can be traced back specifically to the orthodox - comparative, counterfactual, welfarist - concept of harm. I argue that non-counterfactual and non-comparative harm concepts face problems of their own, and don’t fit easily with our best understanding of discrimination; hence they are unsuitable to replace the orthodox concept here. I then propose a non-orthodox - comparative, counterfactual, hybrid (partly welfarist, partly non-welfarist) - concept of harm, which relies on counterfactual comparisons of ways of being treated (rather than well-being levels). I suggest how such a concept can help us handle the problems from counterfactuals, at least for my account of discrimination. I also show that there are similar proposals in other harm-related debates. An upshot of the paper is thus to corroborate the case for a non-orthodox, hybrid concept of harm, which seems better able to fulfil its functional roles in a variety of contexts.

**Keywords** Harm · Discrimination · Wrongness of discrimination · Wrong-making feature · Counterfactual · Welfarist

## 1 Introduction

What is the problem with discrimination? Why is it wrong? If you ask lay people, they might answer that it is wrong, at least in part, because it harms people. It makes them badly off, or

---

✉ Katharina Berndt Rasmussen  
katharina.berndt.rasmussen@iffs.se

<sup>1</sup> Institute for Futures Studies, Stockholm, Sweden

<sup>2</sup> Uppsala University, Uppsala, Sweden

worse off, through no fault of their own; they miss out on opportunities or goods due to some arbitrary feature such as their gender or race. Looking to lawmakers and legal theorists, there appears to be a trend toward a harm-focused approach to the social ills of discrimination that the law seeks to curtail. And researchers in the social and medical sciences also seem predominantly concerned with health- and status-harming features of discrimination.

If you ask philosophers why discrimination is wrong, chances are that they don't mention harm. One might of course argue that moral philosophers who typically understand discrimination in terms of 'disadvantageous treatment' thereby acknowledge harm (welfare-reduction in some sense) as one of the *defining* features of discrimination.<sup>1</sup> But they only rarely consider it as its *wrong-making* feature. That is, definitions of 'discrimination' typically refer to acts, policies or practices that impose relative disadvantages on individuals, due to their social group membership.<sup>2</sup> Yet when it comes to what makes discrimination morally wrong, most proposals concern the *agent's* objectionable motivational states,<sup>3</sup> or objectionable features of the *action* itself.<sup>4</sup> Proposals that do pay attention to the effects of discrimination on the *victim*, mostly do so in deontological or rights-based terms.<sup>5</sup> Very few accounts spell out the wrong-making feature of discrimination in terms of its consequences for the victims' *well-being*. On these latter accounts, typically, there is then nothing distinctively wrong with discrimination, apart from its failure to maximise welfare.<sup>6</sup>

My main question is whether there's a plausible harm-based approach to the moral wrongness of discrimination, that better aligns with the concerns of legal, social, medical theorists and practitioners, and lay people. Such an approach would have the advantage of taking seriously these widespread concerns, making sense of them in a moral-philosophical framework. In section 2, I'll outline these concerns in some more detail, referring to a number of recent studies concerning discrimination. In section 3, I'll introduce a recent account of wrongful discrimination, which divides into (1) a definition of group discrimination, and (2) a characterisation of its moral wrong-making feature in terms of harm.<sup>7</sup> I'll then point out a mismatch between these two components, and improve the account of (1) and (2), reformulating both in terms of *intrapersonal* counterfactual comparisons of the discriminatee's *well-being levels*. Section 4 introduces a number of problems from counterfactuals. These are traced back to the debate on the harm concept, in section 5. Section 6 explores alternative accounts of harm, and shows why they're not useful either. In section 7, I'll propose an alternative revision of (1) and (2), in terms of intrapersonal counterfactual comparisons of *ways of being treated*. I'll suggest that, in order to handle the problems from counterfactuals, 'being treated worse' cannot be analysed in exclusively welfarist terms. I'll also show that there are similar proposals in other harm-related contexts. Section 8 concludes.

<sup>1</sup> Cf. (Altman 2016, para. 1.1).

<sup>2</sup> Cf. (Altman 2016), (Moreau 2010). Some theorists at least implicitly agree with this rough definition, but are mainly concerned with the question when and why such disadvantageous treatment is wrong (Hellman 2008).

<sup>3</sup> These motivational states are characterized as, e.g., taking into account "properly irrelevant" or "trifling" characteristics (Flew 1993, p. 101), as referring to arbitrary facts (Kekes 1995), or as stemming from incorrect prejudice (Dworkin 1985, p. 66), (Ely 1980, p. 153), (Cotter 2006, p. 10).

<sup>4</sup> Such a feature would be e.g. the social meaning of the act, conveying lack of equal concern or respect (Hellman 2008), (Shin 2009).

<sup>5</sup> They concern, e.g., the violation of merit-based rights (Hook 2002), (Goldman 1979), or the violation of entitlements to freedom (Moreau 2010).

<sup>6</sup> (Arneson 1999) and (Lippert-Rasmussen 2014) propose a desert-based prioritarian account of the wrongness of actions, policies and practices in general – which include, but are not limited to, discriminatory actions.

<sup>7</sup> (Lippert-Rasmussen 2014).

## 2 Non-Philosophical, Harm-Related Takes on Discrimination

### 2.1 Legislation

In 1999, in the case of *Davis v. Monroe County Board of Education*, the U.S. Supreme Court ruled that schools which respond with deliberate indifference to student-on-student sexual harassment thereby discriminate on the ground of sex – regardless of whether the school officials acted from sexist motives. According to legal theorist Deborah Brake, this case demarcates a shift in U.S. legal practice toward an improved interpretation of discrimination. This interpretation, she states, “notably and correctly focuses on how institutions cause sex-based harm, rather than on whether officials within those institutions act with a discriminatory intent” (Brake 2001, p. 6). The case thus illustrates a shift towards a harm-based account of legal liability for discrimination.

In a similar vein, but drawing on a larger number of cases and legal texts regarding primarily the jurisdictions of the U.S. and Canada, Denise Réaume states that when discriminatory practices became the object of legal prosecution, a “focus on prejudiced motive seem[ed] to characterize the starting point of both U.S. and Canadian jurisprudence” (Réaume 2002, p. 359, fn. 18). Subsequently, the requirement was changed to negligence of foreseeable harmful effects – or even to just the actual occurrence of such effects. Réaume suggests that this development is due to a more general expansion of the legal concept of fault, and a corresponding expansion of the conception of human interests that the law should protect. She writes that the early “approach constructs discrimination as a cause of action in which malice [a desire to do ill grounded in bigotry or prejudice] is intrinsic to the wrongdoing”, while the current doctrine focuses on the “effect of denying or reducing access to the relevant benefit for those identified by such a characteristic” (Réaume 2002, pp. 360 [358], 373).

In his book-length comparative account of contemporary discrimination law, Tarunabh Khaitan assesses a wealth of data on discrimination law practices in five jurisdictions: Canada, India, South Africa, the United Kingdom, and the United States. Drawing on the emerging “essence of the practice of discrimination law”, he sets out to “discover what it is that discrimination law seeks to do” (Khaitan 2015, p. 119, 118). His answer is that the general purpose is to reduce, and eventually eliminate, significant relative disadvantages between socially salient groups. In the end, he writes, “the concern with relative group disadvantage is ultimately down to a concern for individual well-being” (Khaitan 2015, p. 130).

These studies thus point to a shift in focus, in both current legislation and normative legal theorizing, from malice and intent to the harmful effects of discrimination.

### 2.2 Lay Judgments

Lay judgments are at times at odds with modern legal judgments. This discrepancy is nicely illustrated by a quarrel between the former Archbishop of Canterbury and Lord Justice Law, in 2010. The occasion was a marriage guidance counsellor’s legal appeal against his being sacked on the grounds of discriminating against homosexuals, since he’d refused to provide a gay couple with sex therapy due to religious reasons. In support of his appeal, the former Archbishop stated:

“I wish to dispute that the manifestation of the Christian faith in relation to same sex unions is ‘discriminatory’ [and] that such religious views are equivalent to a person [sic] who is, genuinely, a homophobe and disreputable” (Lord Carey of Clifton 2009: §9).

The judge aptly replied that the former Archbishop had failed to realise that

“the law forbids discriminatory conduct not by reference to the actor’s motives but by reference to the outcome of his or her acts or omissions” (quoted in (Khaitan 2015: 1)).

The conflation of the roles that motives and harmful outcomes play, respectively, for common-sense judgments regarding wrongful discrimination, has been somewhat disentangled by a recent study in psychology. The authors of “The role of intent and harm in judgments of prejudice and discrimination” present a set of experimental studies in which participants observed (or themselves experienced) differential treatment of men and women. The researchers varied information about (a) the agent’s intent to discriminate, and (b) the amount of harm inflicted on the potential discriminatee. They found that the likelihood of judging the observed (or experienced) differential treatment as discrimination co-varied with both variables.<sup>8</sup> When there was doubt about the agent’s intent to discriminate, or clearly no such intent at all, a greater amount of harm correlated with an increase in the participants’ likelihood to judge the differential treatment as discriminatory. When, on the other hand, the agent was clearly acting from prejudiced intent, the participants were more likely to judge the action to be discriminatory, than when there was doubt or no information about intent. In these latter cases, harm was less salient: it only slightly increased the likelihood of judging the action to be discriminatory for those who observed the differential treatment (the increase was greater for those who were subjected to it themselves) (Swim et al. 2003). This study suggests that intent does play a greater role in lay people’s judgments of discrimination than harm does. Importantly for our present purpose, it shows that harm, too, is significant for lay judgments about discrimination.<sup>9</sup>

### 2.3 Medicine, Public Health Science, and Social Epidemiology

Research in medicine and public health studies has examined the harmful effects of discrimination on physical and mental health, as well as life opportunities. The authors of “Racial Discrimination and Blood Pressure” assess the contribution of race discrimination to an as of yet “unexplained twofold greater prevalence of hypertension among the US Black population than among the US White population”. They provide “an approach to documenting how discrimination can harm health”, and indicate that physical harm is one of the problematic features of racist discrimination (Krieger and Sidney 1996: 1370; 1376).

The authors of “Racial and Ethnic Disparities in Mental Health Care for Children and Young Adults” compare the care received by Black, White and Hispanic children and youth for mental health and substance abuse issues. They note that the prevalence of psychiatric

<sup>8</sup> Though the study consistently refers to “judgments of discrimination”, the context (e.g. parallel references to “judgments of prejudice”) suggests that the researchers’ questions to participants and analyses of their answers concern judgments of *wrongful* discrimination. It’s less clear how the participants themselves interpreted the questions.

<sup>9</sup> Cf. “An Examination of Factors Affecting Perception of Workplace Discrimination” (Banerjee 2008), “That’s Not Discrimination: American Beliefs and the Limits of Anti-Discrimination Law” (Eyer 2011, p. 1314), “Pick Your Perspective: Racial Group Membership and Judgments of Intent, Harm, and Discrimination” (Simon et al. 2017).

conditions and substance abuse are roughly similar between these groups. Still, controlling for relevant factors, minority children and youth receive less such care than their White counterparts. The authors further note that the "under-provision of mental health care for minority children contrasts starkly with the high frequency of punitive sanctions [school suspensions, arrest, incarceration] that their behaviors elicit" (Marrast et al. 2016: 10). This suggests that one of the problems with minority discrimination is the two-fold harm of under-provision of resources and over-provision of punishment.

These and other studies also make important contributions to the rapidly emerging field of social epidemiology that, inter alia, investigates the effects of discrimination on individual and population health (in terms of psychological and physiological impairment, but also in effecting harmful behaviours such as substance abuse or unsafe sex-practices). Among the "many pathways [...] by which discrimination could harm health", the field has examined:

- (1) economic and social deprivation;
- (2) excess exposure to toxins, hazards, and pathogens;
- (3) social trauma;
- (4) health-harming responses to discrimination;
- (5) targeted marketing of harmful commodities;
- (6) inadequate medical care; and
- (7) especially (but not only) for Indigenous peoples, ecosystem degradation and alienation from the land. (Krieger 2014: 74, 75)

This list of "pathways" serves to highlight that discrimination can harm individuals and groups both indirectly, mediated by the perception of being discriminated against, and directly, as an inherent feature of the treatment. Conditions (1), (2), and (6) clearly point to racially differentiated economic, social, geographical, and medical patterns of risk and harm. These patterns are themselves instantiations of the *intrinsic* features of wrongful discrimination. Conditions (3) and (4), in contrast, refer to the (collective and individual) responses to the *perception* of being treated differently, that is, of being on the losing side of these patterns.

This sketch of harm-related concerns of non-philosophers who work on and think about issues of discrimination will here provide the background for my quest for a harm-based approach to the wrongness of discrimination within moral philosophy. I'll now take a closer look at Kasper Lippert-Rasmussen's (2014) proposal for a distinct and intrinsic wrong-making feature of discrimination in terms of harm.

### 3 A Harm-Based Account of the Wrongness of Discrimination

Lippert-Rasmussen defines group discrimination as follows.

**(GD)** "An agent, X, [group] discriminates against someone, Y, in relation to another, Z, by  $\phi$ -ing (e.g., hiring Z rather than Y) if, and only if:

- (i) There is a property, P, such that Y has P or X believes that Y has P, and Z does not have P or X believes that Z does not have P,
- (ii) X treats Y worse than he treats or would treat Z by  $\phi$ -ing, [...]

- (iii) It is because (X believes that) Y has P and (X believes that) Z does not have P that X treats Y worse than Z by  $\phi$ -ing” and
- (iv) “P is the property of being a member of a certain socially salient group (to which Z does not belong)”. (Lippert-Rasmussen 2014, p. 15; 26)<sup>10</sup>

(GD) is then further qualified by an analysis of ‘because’ in condition (iii):

**(B)** “X [ $\phi$ -es] because (X believes that) Y has P and (X believes that) Z does not have P if, and only if, (i) the [belief] that Y, and not Z, has P is part of X’s [...] motivating reason for  $\phi$ -ing, or (ii) the fact that Y, and not Z, has P causally explains X’s  $\phi$ -ing and this in turn is causally explained by the fact that people with P are often treated worse than those without P in the sense given by (i).” (Lippert-Rasmussen 2014, p. 38)<sup>11</sup>

(B) ensures that both (i) direct and (ii) indirect discrimination are covered by the analysis.<sup>12</sup> Furthermore, the concept of social salience is analysed as follows:

**(S)** “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.” (Lippert-Rasmussen 2014, p. 30)<sup>13</sup>

Next, the account states when and why discrimination, thus defined, is morally wrong. Lippert-Rasmussen starts out with the goal of accounting for the *distinctive* and *intrinsic* wrongness of discrimination in terms of harm. Though he subsequently abandons these two desiderata, I will here try to push a bit harder in adhering to them:

**(D1)** Intrinsicness: discrimination is wrong as such, i.e. “it is wrong in virtue of properties that are true of discriminatory acts by definition.” (Lippert-Rasmussen 2014, p. 104)

**(D2)** Distinctiveness: “there is at least one feature common to all forms of wrongful discrimination by virtue of which they are morally wrong.” (Lippert-Rasmussen 2014, p. 105)

Lippert-Rasmussen’s claim for the harm-based wrong-maker of discrimination states:

**(HW)** “an instance of discrimination is [prima facie] wrong, when it is, because it makes people worse off, i.e., they are worse off given the presence of discrimination than they

<sup>10</sup> Lippert-Rasmussen suggests an additional condition: (v) “ $\phi$  is a relevant type of act, policy, or practice, and there are many acts etc. of this type, and this fact makes people with P (or some subgroup of these people) worse off relative to others, or  $\phi$  is a relevant type of act etc., and many acts etc. of this type would make people with P worse off relative to others, or X’s  $\phi$ -ing is motivated by animosity towards or dislike of individuals with P or by the belief that individuals who have P are inferior or ought not to intermingle with others.” (Lippert-Rasmussen 2014, p. 28). I’ll ignore this cumbersome (and to my mind non-necessary condition). Those who object may easily add it onto my final proposal below.

<sup>11</sup> The original formulation of (i) is: “the *thought* that Y, and not Z, has P is part of X’s *direct* motivating reason for  $\phi$ -ing”, (Lippert-Rasmussen 2014, p. 38, italics added). My revised formulation has been accepted by Lippert-Rasmussen (personal communication, March 2016).

<sup>12</sup> I’ll neglect further specifications of the divide, as outlined by (Lippert-Rasmussen 2014, chap. 2).

<sup>13</sup> While Lippert-Rasmussen states the condition as merely sufficient, it should be understood as both sufficient and necessary (personal communication, March 2016). Note that, on pain of rendering the definition circular, the “structure of social interactions” mustn’t be *analysed* by reference to discrimination. This is of course compatible with this structure (independently analysed) being *caused* by phenomena, some of which we would want to label discrimination in their own right.

would have been in some suitable alternative situation in which the relevant instance of discrimination had not taken place.” (Lippert-Rasmussen 2014, p. 154f)<sup>14</sup>

Let’s take a closer look at (GD) and (HW). Clearly, they concern two distinct phenomena: (GD) concerns what *constitutes* discrimination itself, (HW) concerns the moral relevance of its harmful *effects*.<sup>15</sup> This is quite alright on its own terms. But by the lights of the two desiderata, some question marks are raised. First, (D1) calls for an *intrinsic* wrong-making feature. Yet (HW) does not pick out any of the properties referred to by (GD). Second, considering (D2), it remains an open question whether any action that falls under (GD), yet doesn’t satisfy (HW), would *not* be a wrongful form of discrimination. Which is to say, it remains an open question whether (HW) manages to pick out a feature that is “common to *all* forms of wrongful discrimination”. Third, considering that the proposal is framed as a harm-based account of discrimination, it’s odd that ‘harm’ is mentioned neither in (GD) nor (HW).

I’ll attempt to straighten out the first and third question mark in the remainder of this section. The second question mark, regarding (D2), is the main topic for the rest of this paper.

Concerning (D1), there’s a *double* mismatch between the features that are picked out by (GD) and (HW), respectively. (HW) states that “[people] are *worse off* given the presence of [an instance of] discrimination than *they* would have been in some suitable alternative situation”. Thus (HW) relies on an *intrapersonal* comparison between the discriminatee’s actual and counterfactual states, in terms of the discriminatee’s levels of *well-being* in these states.

On the other hand, group discrimination itself is claimed to be “*essentially comparative with respect to individuals*” (Lippert-Rasmussen 2014, p. 16). (GD) requires that “*X treats Y worse than he treats or would treat Z by  $\phi$ -ing*”. Thus, (GD) relies on *interpersonal* comparisons between the potential discriminatee Y and another individual Z. Moreover, the comparison doesn’t concern their levels of well-being, but rather their ways of *being treated*.

The double mismatch thus concerns (a) intrapersonal vs. interpersonal comparisons of (b) levels of well-being vs. kinds of treatment. Is there a way to tweak the account to provide us with a harm-based wrong-making feature that is *intrinsic* to discrimination, as requested by (D1)?

Here’s one proposal: we restate (GD) in intrapersonally comparative terms, to avoid mismatch (a), and in terms of levels of well-being, thereby avoiding mismatch (b).

**(GD1)** An agent, X, group discriminates against someone, Y, by  $\phi$ -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  $\phi$ -ing, X makes Y worse off than Y would have been, had Y not had P or had X not believed Y to have P,
- (iii) It is because (X believes that) Y has P that X makes Y worse off by  $\phi$ -ing, and
- (iv) P is the property of being a member of a socially salient group.

<sup>14</sup> I’ve added the ‘prima facie’ caveat to clarify that a discriminating action’s harmfulness “counts as a reason in favor of its being impermissible to perform”, yet may be outweighed by other reasons (Lippert-Rasmussen 2014, p. 103).

<sup>15</sup> Of course, discriminatory acts can be wrong because of a number of extrinsic features: they may (or may not) involve breaking promises, violating rights, failing to maximise total utility, and the like; cf. (Lippert-Rasmussen 2014, p. 104). But here we are looking for an *intrinsically* wrong-making feature that makes such acts wrong even if they honour all promises and rights, maximise overall utility, etc.



(GD1) makes an intrapersonal welfare level comparison between actual Y-with-P and counterfactual Y-without-P. It thus fits nicely with (HW). I'll only slightly revise (HW), simply to make it more precise:

**(HW1)** An instance of group discrimination,  $\phi$ , by X against Y is (prima facie) wrong, when it is, because it makes Y worse off than Y would have been, in some suitable alternative situation in which the relevant instance of discrimination had not taken place.

Some may object that this version of Lippert-Rasmussen's account is problematic. One concern is that it seems to make group discrimination morally wrong by definition.

### 3.1 The Moralised-Definition Objection

The objection is that the question of moral wrongness shouldn't be settled per definition, and that we should allow for morally innocuous instances of group discrimination.

I agree with the first conjunct of this objection but maintain that (GD1) satisfies it. It doesn't provide a moralised definition. While it's formulated in evaluative terms of 'making worse off', it doesn't make any moral claims. Such a claim is introduced by (HW1), a separate and non-definitional claim pertaining to the wrong-maker of discrimination. The definition is thus non-moralised. The entire account of (GD1) and (HW1), of course, isn't – and shouldn't be.

Since (HW1) picks out features of (GD1), as demanded by D1, the account implies that all instances of group discrimination are prima facie morally wrong. This is the complaint of the second conjunct of the objection. I do, however, question its reasonableness. (GD1) carves out a very specific social phenomenon (*group* discrimination), which most of us take issue with. (HW1) then merely spells out why we are (morally justified in) taking such issue with it.

### 3.2 The Too-Revisionist Objection

Another objection is that restating (GD) and (HW) in intrapersonally comparative terms is too revisionist, considering that interpersonal comparisons often figure in everyday assessments of discrimination, as well as in many theories of discrimination.<sup>16</sup> But restating the account in intrapersonal terms not only helps avoid one of the two internal mismatches, but also mitigates problems of the following kind. Imagine:

**Robinson Crusoe**, stranded on his island, meets Friday. They keep each other company, but Robinson makes Friday quite miserable: he makes him work unnecessarily hard, for a very small share of the available food, he curses at him, etc. The reason is that Friday is black, and Robinson holds blacks to be inferior to whites like himself.

Does Robinson's behaviour become wrongful discrimination only when *another*, non-black individual, call her Saturday, is introduced to the island (or could be introduced, alongside Friday) and treated better? No. Robinson's behaviour is discrimination all along, if it's the case that, *had Friday not been black*, Crusoe would have treated *him* better. This suggests that what's relevant is Friday's property of being black and its counterfactual absence, rather than the presence of some third party who lacks this property. (GD1) captures this insight.<sup>17</sup>

<sup>16</sup> Cf. (Altman 2016, para. 1.1).

<sup>17</sup> Lippert-Rasmussen has accepted a reformulation along these lines (personal communication, March 2016).



The reason why interpersonal comparisons often do play a role in everyday assessments, as well as philosophical accounts of discrimination is, I want to suggest, of pragmatic origin. My present account can acknowledge this pragmatic reason, without taking it to define discrimination. (GD1) compares the well-being resulting from X's  $\phi$ -ing for actual-Y-with-P, with how counterfactual-Y-without-P would have fared. However, in order to *determine* how Y counterfactually would have fared, it may be useful to consider evidence as to how X actually behaves toward others – such as Z-without-P. We might realise only *after* observing Robinson interact with Saturday, that he would have made Friday much better off, had Friday not been black.

Interpersonal comparisons can thus be pragmatically important, by providing us with evidence for determining how X would have acted in the relevant counterfactual situation. Acknowledging this pragmatic role helps explain why many theories of discrimination rely on comparisons between persons: they pick up on our practice of doing just that when we assess whether someone discriminated against another or not.<sup>18</sup> My upshot is that this practice needn't be taken to imply that interpersonal comparison is a core feature of discrimination.

A further upshot: this makes the account more clearly *harm*-based. It relies on intrapersonal, counterfactual comparisons of well-being levels, reminiscent of the orthodox harm concept in comparative, counterfactual, welfarist terms.<sup>19</sup> If we accept that harm is best understood in these terms, then we see the revised account of (GD1) and (HW1) as essentially harm-based. This straightens out the above second question mark.

To fully understand the revised account and apply it, we need to know what exactly 'some suitable alternative situation' in (HW1) refers to. It should be the same counterfactual state that (GD1) refers to in clause (ii): 'had Y not had P or had X not believed Y to have P'. In the next section, I consider three proposals by Lippert-Rasmussen and show them to be problematic. I then connect their problems to a general problem with 'harm'.

#### 4 The Relevant Counterfactual

The *Straightforward Account* states that "we hold everything else other than the discriminatory act and what follows causally downstream from it constant" (Lippert-Rasmussen 2014, p. 157). But this won't do, Lippert-Rasmussen concedes: consider a Jewish job applicant who is not hired because she is Jewish, who therefore leaves 1939 Germany, and thereby narrowly escapes deportation. Surely, she is better off than she would have been had she been hired and – down the causal stream – stayed on, suffered deportation, and worse. Still, the refusal to hire her, due to her being Jewish, is clearly a paradigmatic case of wrongful group discrimination.

Lippert-Rasmussen proposes the *No Discrimination Baseline Account*: "we should compare the actual outcome to the one that would have obtained if the discriminatory act had not been performed and no [sic] else had performed discriminatory acts in the future" (Lippert-Rasmussen 2014, p. 158). We should thus disregard the counterfactual where the applicant, let's call her Sasha, upon being hired had been deported, since the deportation itself would have been an instance of discrimination.

<sup>18</sup> This pragmatic value might make interpersonal comparisons essential for a legal concept of discrimination. Given their different functional roles, such a divergence between legal and moral concepts of discrimination shouldn't be all too surprising, nor indeed, worrisome.

<sup>19</sup> See e.g. (Bradley 2012, p. 396), (Feinberg 1987, p. 33 f.), (Parfit 1986, p. 69), (Kagan 1991, p. 84), (Norcross 2004, p. 150), (Bradley 2009, p. 65).

Yet surely, it should be objected, this still isn't good enough. This account would prevent us from calling the following non-hiring decision wrongful discrimination.

**Sasha** is not hired for a job in 1939 Germany because she is Jewish. This causes her to emigrate. Had she been hired and stayed on, she would have been killed by an (entirely non-discriminatory) stray bullet from a botched bank robbery.

We surely shouldn't conclude that the non-hiring decision was then not wrongful discrimination after all.

A third, *Ideal Baseline Account*, fares no better. It states the relevant counterfactual in terms of "had everyone else acted morally permissible from the time of the discriminatory act and henceforth" (Lippert-Rasmussen 2014, p. 158, fn. 9). To see why this won't do, consider the following case.

**Charlie**, who's unemployed, needs to apply for jobs to be eligible for unemployment benefits. She knows there is widespread sexism in the work sector. Desiring to not get disappointed (and to keep her benefits), Charlie applies for a job without really making an effort. She's rejected – because of her gender.

The suitable alternative situation is one where everyone else (including the employer, I surmise) acts morally permissible from the time of the hiring decision and henceforth. Since we should hold everything else constant, including Charlie's desires and beliefs, she would have applied for the same job, with the same lack of effort. Arguably, she then still wouldn't be hired – this time because of her poor application. Still, the actual, gender-based non-hiring decision seems wrongfully discriminatory. To deal with such cases of socially pervasive and individually internalised sexism, it seems one would have to exclude even (a sufficient amount of) past discrimination. But then it becomes increasingly indeterminate what could be held constant: what would Charlie's beliefs and desires have been in such a world? What would her options be? Would she even have existed?<sup>20</sup>

These cases spell trouble, considering desideratum (D2). They are cases where, though no one is made worse off (compared to the relevant counterfactual), we still want to claim wrongful discrimination. By generating false negatives, they show that our current account fails to identify the common wrong-making feature shared by all cases of wrongful discrimination.

Moreover, these are more than just quick counter-examples. They point to a more general and pervasive problem: accounts of discrimination relying on counterfactual comparisons of well-being inherit many of the problems known from the debate on the harm concept.

## 5 The Orthodox Account of Harm

The above counterexamples are problems of *pre-emption*: things might have gone even worse for the potential victim, had the event in question not occurred. Compare this classic harm-problem:

<sup>20</sup> Similar problems, concerning the appropriate baseline for comparison, arise for accounts of reparation or rectification of past wrongs, as discussed by e.g. (Waldron 1992) and (Morris 1984). I want to thank an anonymous referee for bringing the close connection of these debates to my attention.

“Suppose Batman drops dead of a heart attack. A millisecond after his death, his body is hit by a flaming cannonball. The cannonball would have killed Batman if he had still been alive. So the counterfactual account entails that the heart attack was not harmful to Batman. It didn’t make things go worse for him. But intuitively, the heart attack was harmful. The fact that he would have been harmed by the flaming cannonball anyway does not seem relevant to whether the heart attack was actually harmful.” (Bradley 2012, p. 397)

In pre-emption cases, counterfactual accounts of harm risk falsely (counter-intuitively) classifying the event as not harmful. As we’ve just seen, our account of discrimination, relying on such an account of harm, runs the analogous risk of missing intuitively clear-cut cases of wrongful discrimination.

Then there’s the related problem of *over-determination*. Consider another classic:

“X and Y shoot and kill me. Either shot, by itself, would have killed.” (Parfit 1986, p. 70)

Since things wouldn’t have been better for the victim, had X’s shot not occurred, it counter-intuitively follows that X didn’t harm the victim. The same of course holds for Y’s shot. Yet it’s counterintuitive to say that neither X nor Y harmed me.<sup>21</sup> An analogous case for our account of discrimination is the following.

**Elliot** applies for a job. Her application is rejected by two members of a hiring committee, one of whom holds disabled people to be incompetent, and has observed Elliot using a wheelchair. The other, however, skipped lunch, and rejects everyone due to low blood sugar.

If unanimous acceptance is required for hiring decisions, Elliot wouldn’t have gotten the job even if she had not been (believed to be) disabled. Due to the presence of the hungry committee member, our account fails to mark the ableist committee member’s rejection as an instance of wrongful discrimination, on any of the suggested counterfactual baselines. This is clearly inadequate.

In the harm-debate, there are further cases of false negatives, e.g. stemming from the *non-identity problem*. Consider this well-known case:

“*The 14-Year-Old Girl*. This girl chooses to have a child. Because she is so young, she gives her child a bad start in life. Though this will have bad effects throughout this child’s life, his life will, predictably, be worth living. If this girl had waited for several years, she would have had a different child, to whom she would have given a better start in life. [...] We cannot claim that this girl’s decision was worse for her child.” (Parfit 1986, p. 358f.)

To refocus on discrimination, imagine:

**EU-policy** (A) will considerably improve the lives of one million existing Europeans who are quite well off, but as a consequence will result in ten million barely-worth-living African lives. The alternative policy (B) would only slightly improve these existing one million European lives, but instead result in ten million quite well off African lives, though belonging to different people.

<sup>21</sup> Parfit bites the bullet that, individually, X and Y don’t harm me; he holds that they *together* harm me.

The decision to implement policy (A), due to a disregard for African lives, wouldn't qualify as wrongful group discrimination: these ten million wouldn't have been better off in the suitable alternative situation. Rather, they wouldn't have existed at all. Again, our account seems to be missing intuitively clear-cut cases of wrongful discrimination.

On second thought, this seems an easier bullet to bite, given some qualifications. It would be justified for the Africans, with lives barely worth living, to complain about EU-policy (A). But would it be correct for any one of them to claim that it wrongfully discriminated against *her*? I want to suggest that the complaint plausibly concerns the wrongful discrimination against her *birth cohort*, since the policy made it worse off than *it* (containing the same or other members, and on some aggregative view) would have been under policy (B). This exemplifies an argumentative strategy for handling non-identity cases in the present context.<sup>22</sup>

There is a fourth class of objections which are usually brought against comparative counterfactual harm concepts: cases of *omission*. Consider a further classic:

“Suppose Batman purchases a set of golf clubs with the intention of giving them to Robin, which would have made Robin happy. Batman tells the Joker about his intentions. The Joker says to Batman, ‘why not keep them for yourself?’ Batman is persuaded. He keeps the golf clubs. The comparative account entails that Batman has harmed Robin, because Robin would have been better off if Batman had not kept the clubs. But it seems implausible to say that Batman has harmed Robin.” (Bradley 2012, p. 397)

Thus, comparative counterfactual harm concepts also generate false positives. To see how the discrimination-account dodges these cases, consider the adapted Batman-case:

**Batman**, upon buying the golf clubs as a gift to Robin, learns from the Joker that Robin's in fact transgender. Because of this fact, he decides to keep the clubs for himself.

Now it does seem fitting to say both that Batman discriminates against Robin, qua member of the socially salient trans community, and that this is *prima facie* wrong (even though it may be overall permissible, as Batman has no obligation to give gifts to Robin). The decision is an instance of *wrongful* discrimination because it makes Robin worse off than they could have been. But it's *discrimination*, to start with, just in case it's due to a socially salient property: being transgender.

To compare, in the original case, Batman makes the same decision – for reasons concerning selfishness, rather than any socially salient property. Since thus (GD1)'s condition (iv) isn't satisfied, there's no group discrimination, and hence there cannot be *wrongful* group discrimination. The lexical structure of the account – an action needs to satisfy (GD1), for the disadvantage it causes to be given moral relevance by (HW1) – allows it to correctly label acts of omission as wrongful discrimination only if they are instances of discrimination, and to be silent on all other cases.

In handling non-identity and omission cases, my harm-based account of discrimination demonstrates an interesting resilience to some of the problems that usually trouble comparative counterfactual welfarist accounts of harm. Still, pre-emption and over-determination remain problematic. Within the harm-debate, alternative accounts have been suggested to avoid these troubles. Let's see whether they hold any promise.

<sup>22</sup> It also exemplifies an application of the discrimination-account to collectives.

## 6 Alternative Harm Concepts

### 6.1 Amended Counterfactual Accounts

Judith Jarvis Thomson suggests a *Revised Counterfactual Comparative Account*: the relevant counterfactual state is the one that the agent prevents “by the same means” as by which she brings about the actual state (Thomson 2011, p. 448). The idea is to fix the relevant counterfactual on some given or obvious trajectory of events, actual deviation from which is caused by the agent’s action. Yet this account runs into trouble when there is no obvious trajectory, or when such a trajectory isn’t prevented *by the agent*, e.g. due to over-determination or preemption. In Elliot’s case, ‘being hired’ seems to be on the “obvious” trajectory, yet isn’t prevented by the ableist committee member’s rejection, since it hinges as much on another (hungry) committee member’s rejection. In Sasha’s case, moreover, it’s hard to know what goes on the “obvious” trajectory: ‘being hired’ should be on it, but shouldn’t ‘being deported the very next day’, too? If this is denied, since Sasha would have been killed before that by the stray bullet, should the stray bullet go on the trajectory instead? This seems rather non-obvious.

Nathan Hanna proposes that the relevant counterfactual state is the closest possible world in which the agent doesn’t act with the same “fundamental intention” as in the actual world (Hanna 2016, p. 18). Yet again, in Sasha’s case, the closest possible world in which the employer doesn’t have any anti-Semitic intentions, and thus hires Sasha, is the one in which Sasha steps in the way of a stray bullet.<sup>23</sup>

Generally speaking, it seems that for any counterexample there is some way to revise the counterfactual approach to handle it.<sup>24</sup> But for any such revised counterfactual account, some new counterexample can be devised. We’re getting nowhere. In the light of this, it’s worth exploring: could non-comparative or non-counterfactual accounts of harm do a better job?

### 6.2 Non-Comparative Accounts

One option is to characterize states of harm non-comparatively: in terms of death, illness, pain, disfigurement, and similar items on some list; or in more unified terms of intrinsic badness.<sup>25</sup> These accounts face many problems of their own.<sup>26</sup> There’s, moreover, a decisive objection against using them in the discrimination context. Discrimination *may* of course cause the victim’s pain, disfigurement, etc., or put her in an intrinsically bad state. But it needn’t. And typically, it doesn’t. The distinctive problem with discrimination is that the victims are made worse off than they *would have been*. A rich and famous – well off – female movie star is still wrongfully discriminated against if her pay-check is lower than it would have been, had she been male (as may be evidenced by her male co-star’s actual pay-check).

In brief, non-comparative concepts of harm aren’t useful here since discrimination is an inherently comparative notion. I’ll therefore accept a comparative concept. Could the *counterfactual* feature be replaced?

<sup>23</sup> Analysing the wrongness of discrimination in terms of intentions would also spell trouble for non-intentional discrimination.

<sup>24</sup> Other counterfactual accounts target other problems, e.g. the non-identity problem (Gardner 2015). (Klocksiesm 2012) also targets preemption and over-determination cases; for criticism see fn. 24 below.

<sup>25</sup> See (Harman 2009), (Shiffrin 1999).

<sup>26</sup> See (Thomson 2011), (Bradley 2012), (Hanser 2008), (Holtug 2002), (Northcott 2015).

### 6.3 Non-Counterfactual Accounts

This means replacing the counterfactual baseline with some other comparative baseline. *Salient-Contrast Comparative Accounts* compare the actual course of events with some salient contrast alternative, e.g., the alternative on which “the participants in a conversational context consciously focus” (Norcross 2004, p. 173, fn. 13).<sup>27</sup> This implies that altering the salient contrast alternative may suffice to change our judgment on whether someone is harmed. Norcross illustrates his view with the case of Bobby Knight, who violently chokes his victim – intuitively harming her – but who would otherwise “have torn both her arms from her body and beaten her over the head with them” (Norcross 2004, p. 166).

For starters, it seems highly doubtful that anyone would concede that Bobby Knight’s choking didn’t harm the victim – even when it’s entirely clear from the conversational context that the salient alternative was Bobby Knight’s dismembering the person. The choking harms the person. The dismembering would have harmed her even more. That doesn’t make the choking non-harmful.<sup>28</sup> Furthermore, as I have argued, when discussing discrimination, the conversational context focuses on the counterfactual state of Y-not-having-P as the salient alternative. This account would thus collapse into a counterfactual account.

*Temporal Comparative Accounts* instead compare the victim’s resulting state with her state prior to the event.<sup>29</sup> But the underpaid female movie star may be discriminated against even if the pay check makes her better off than she was before. One could easily come up with further non-counterfactual baseline proposals. The trouble with all these accounts seems, generally, that one as easily can come up with counterexamples for any such baseline.

Other proposals focus on *events* rather than states. According to the *Event-Based Account*, Y suffers harm if and only if Y is subject to a level-1 harm: an event of losing a basic good; or to a higher-level harm: an event of being prevented from acquiring a basic good, or of being prevented from being prevented from a lower-level harm.<sup>30</sup> This metaphysical manoeuvre should, however, not distract us from examining the basic building blocks: the event of *losing* something is an event of going from having it (prior) to not having it (post). The event of *being prevented* from acquiring something is most naturally understood as an event interfering with some trajectory of other events that (counterfactually) would have resulted in its acquisition.<sup>31</sup> This now looks like a *Double Comparative Account*, inviting counterexamples against both counterfactual *and* temporal comparative accounts. Consider Elliot and Sasha: they don’t lose a job, since they never had it. And they are not prevented, by the rejecting agent, from being better off, since something else (the hungry committee member, or the stray bullet) would have made sure they never would be. I conclude that, as of now, no intuitively plausible and useful replacement for the counterfactual component has been forthcoming.

This paper started out from an outline of discrimination and its intrinsic and distinctive wrong-making feature in counterfactual, comparative, and welfarist terms. This invoked not only the orthodox harm concept, but also invited many of its problems. It now appears that it

<sup>27</sup> (Norcross 2004, p. 167 f.) proposes that an event is harmful relative to a salient contrast event; (Northcott 2015), in a similar vein, refers to a salient contrast-cause and contrast-effect.

<sup>28</sup> Cf. (Bradley 2012, p. 407). Similar criticism can also be brought against (Klocksiesm 2012) and his *Contextual Counterfactual Account*.

<sup>29</sup> See e.g. (Thomson 2011, p. 443). For criticism, see (Thomson 2011, pp. 443–446).

<sup>30</sup> Cf. (Hanser 2008). For further criticism of his account, see (Thomson 2011, pp. 455–458).

<sup>31</sup> Cf. (Klocksiesm 2012, p. 288), who also notes that this double-comparative nature makes Hanser’s account a non-unified, and unnecessarily complicated one.

would be fruitless to try to adapt the discrimination-account to non-comparative, or non-counterfactual harm concepts. These are riddled by problems of their own, and just invite new problems into our discrimination-account.

This brings the welfarist component of the orthodox harm concept into focus. Could a non-welfarist harm concept come to the rescue? In the next section I'll explore a cue in Lippert-Rasmussen's definition of discrimination that points into this direction, and provide a rough sketch of how this might help.

## 7 Acknowledging an Ambiguity

Let's recapitulate my two moves to match (GD) and (HW), to satisfy desideratum (D1). First, I reformulated (GD) in terms of *intrapersonal* comparisons (for which I also gave an independent reason). Second, I replaced 'being treated worse' in (GD) with 'being made worse off'. Let's backtrack halfway and try an alternative second move, reformulating both (GD1) *and* (HW1) in terms of 'being treated worse'.

**(GD2)** An agent, X, group discriminates against someone, Y, by  $\phi$ -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  $\phi$ -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 (X believes that) Y has P that X *treats Y worse* by  $\phi$ -ing, and
- (iv) P is the property of being a member of a socially salient group.

To see what this implies, consider again Sasha, who was not hired because she was Jewish. She therefore avoided the stray bullet and was thus *made better off*. Still, there's a sense in which she was *being treated worse* than she would have been, had she not been Jewish. Consider the other non-hiring cases. Charlie and Elliot would still not have been hired, and thus not better off, in the relevant counterfactual state where they hadn't been female or disabled, respectively. Still, there's a sense that they were *treated worse* than they would have been otherwise.

I want to suggest that this expression involves an important ambiguity: 'Y is being treated worse by X' can refer to the relative *outcome* of X's actions *for Y*: X makes Y worse off. I'll call this the *welfarist* meaning, which has been my main focus so far. But 'Y is being treated worse by X' can also refer to the *kind of action* X performs: X treats Y as inferior, less worthy of consideration, less capable, or of lesser moral standing. This is a – rough and incomplete – sketch of what I'll call the *non-welfarist* meaning of 'being treated worse'.

I suspect that part of the intuitive plausibility of Lippert-Rasmussen's account of discrimination hangs on this ambiguity, and specifically the non-welfarist meaning it conveys.<sup>32</sup> Focusing exclusively on the welfarist meaning, as we did in examining outcomes in orthodox

<sup>32</sup> (Lippert-Rasmussen 2014, p. 18) writes: "Disadvantageous treatment should be distinguished from treatment that causes harm", and thus leaves open the possibility of a non-welfarist sense of disadvantageous treatment, although he doesn't develop this idea further.



terms of harm, not only invites a range of problems. It also neglects a feature that made the account sound intuitively plausible in the first place.

Consider an extreme case, where an intuitively wrongful discriminating action directly causes the victim to be *better off*.

**Alex** is a philosopher. A colleague of hers has to submit his logic paper by Monday and urgently needs a proof-reader. He asks all his male fellow philosophers (none of whom has the time), but not Alex. Since Alex is a woman, he cannot imagine her to be good at logic. (Had he only read some of her articles, he would have realised his mistake.) Alex can spend the weekend at the beach, which makes her better off than she would have been, had she dutifully helped her colleague with the tedious proof-reading at her office desk.

A bystander – even Alex herself – could reasonably complain about the colleague’s sexism (gender-based discrimination). The colleague couldn’t invalidate the complaint by pointing out that Alex benefited from being off proof-reading duty. (GD2) helps us make sense of these claims, once we recognise the double meaning of ‘being treated worse’: Alex is treated worse in the non-welfarist sense, as inferior, simply on grounds of her gender.

I propose that this feature is part of what makes group discrimination wrong. It should therefore be stated explicitly in the claim concerning the wrongness of discrimination.

**(HW2)** An instance of group discrimination,  $\phi$ , by X against Y, on grounds of P, is (prima facie) wrong, when it is, because X by  $\phi$ -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 X not believed Y to have P.

The sense of (b) is captured by many rival accounts of the wrongness of discrimination, which state that discrimination distinctively *wrongs* its victims, or constitutes *a wrong* to them (over and above possibly being wrong for other reasons).<sup>33</sup> The common denominator for these accounts is an appeal to equality: e.g., discrimination is wrong in virtue of failing to pay “recognition respect for someone’s standing as [an equally valuable and autonomous] person” (Eidelson 2015, pp. 74–75); in virtue of violating equal entitlements to certain freedoms (Moreau 2010); of expressing “the unequal humanity of the other” (Hellman 2008, p. 38)<sup>34</sup>; of acting on an attitude that some are not entitled to equal respect (Dworkin 1985); or in virtue of being “premised on the belief that some types of people are morally worthier than others” in the sense of meriting greater moral concern (Alexander 1992, p. 161). By spelling out ‘treating worse’ as ‘making worse off *or* treating as inferior’, the revised account accommodates this sense. This helps it deal with the above-mentioned challenges.

Does this mean that the welfarist component of ‘harm’ becomes redundant? I’ll deal with this worry now.

<sup>33</sup> Cf. (Altman 2016, para. 4.1), (Moreau 2010, p. 145 f.), (Hellman 2008, p. 15 f.). (Lippert-Rasmussen 2014, pp. 103–152) presents objections to a number of these, considered as free-standing accounts. However, referencing these accounts here merely serves to point to a central concern that they – in contrast to harm accounts – aim to capture.

<sup>34</sup> Hellman takes a slight detour in analysing the wrongness of discrimination in terms of being “demeaning”, which in turn is analysed as requiring “an expression of the unequal humanity of the other” (as well as a certain status of the agent). I’m here cutting through to the account’s core. Cf. (Lippert-Rasmussen 2014, p. 134) for a similar analysis.

## 7.1 The Redundancy Objection

I've argued that we need a non-welfarist component of 'being treated worse' in order to avoid false negatives, as in the cases of Sasha, Charlie, Elliot, and Alex. But do we really need the welfarist component now? That is, are there cases of wrongful discrimination where no one is treated as inferior, such that the welfarist component becomes decisive? Consider two cases.

**Kim** is new at her workplace and turns out to be the only woman around. Her male co-workers treat her in a highly respectful and professional way. They couldn't think of asking her to join them at their simple lunch joint nearby, or at their raucous after-work place. As a result, Kim becomes isolated and unhappy at work.

**A local charity** regularly organises soup kitchens for the area's homeless. They operate on very low funding, and make the soup from food donations from the local supermarket. The latest donation allows them to cook a nourishing pork stew. When they hand it out to the homeless, those with Muslim background decline and go hungry, in contrast to their non-Muslim fellows.

In neither case are people treated as inferior. Still, they're made worse off, due to their gender and religion, respectively. The welfarist component of my account allows us to characterise these cases as instances of wrongful discrimination.

In holding on to a welfarist *and* a non-welfarist component, my account is similar to Thomas Scanlon's idea that discriminatory acts are "wrong because of their consequences – the exclusion of some people from important opportunities – and because of their meaning – the judgment of inferiority they express and thereby help to maintain." (Scanlon 2008, p. 73) Opportunity loss is a specific form of welfare reduction; demeaning expressions are one way of treating people as inferior. My account is generic in the sense that its two components can – but needn't – be interpreted along Scanlonian lines. It's allows for other currencies of welfare, and other requirements of equal treatment, to be plugged into clauses (a) and (b), respectively.<sup>35</sup>

## 7.2 The no-Harm Concept Objection

Another worry is that the account is no longer a harm-based, but rather a hybrid, harm- *and* inferior treatment-based account. Doesn't admitting that we cannot do without a non-welfarist component amount to admitting that the harm concept just cannot fulfil the role we're trying to assign it here?

This is certainly the case if we presuppose that harm is necessarily – and exclusively – a welfarist notion. But this presupposition has been questioned even in other contexts. The idea that harm has a non-welfarist dimension is not new. Some mention it only to reject it (Bradley 2012, p. 402), (Holtug 2002, pp. 377, 382). However, the need to explicitly reject the idea suggests that 'harm' does have this double connotation in everyday talk and thought.

Others point out the usefulness of this idea. Woodward analyses harm in the context of the Non-Identity Problem, and emphasises a non-consequentialist dimension, over and above a consequentialist, welfarist dimension (Woodward 1986, p. 818). Feinberg analyses 'harm'

<sup>35</sup> A more comprehensive analysis of the similarities and differences between Scanlon's and my accounts merits an article of its own. Here, it must suffice to note that my attempt to develop a plausible harm-based account of the wrongness of discrimination turns out to have moved it in a Scanlonian direction.

within a legal context and carves out two distinct but non-exclusive senses: (1) the set-back of an individual's interests (corresponding to my welfarist dimension), and (2) the wronging, or unjust treatment, of the individual (my non-welfarist dimension) (Feinberg 1987, p. 36). There are differences between these proposals and my own – both in content and in context. But clearly, the debate on a purely welfarist sense of harm has not been settled.

## 8 Conclusions

I've made a case for a hybrid – welfarist and non-welfarist – harm concept, which fulfils its functional role as a wrong-making feature of group discrimination and captures some of our lay and expert judgments in this context. It turns out that this concept not only partly aligns my account of the wrongness of group discrimination with rival accounts. It also in itself aligns closely with harm concepts that have proven useful in other contexts. Most importantly, it helps us deal with conundrums from counterfactuals that have been shown to befall exclusively welfarist harm-based discrimination-accounts.

**Acknowledgements** This article is a result of the research project “Harm: The Concept and Its Relevance” at Uppsala University, generously funded by The Swedish Foundation for Humanities and Social Sciences (RJ; P14-0212:1). It was written at the Institute for Futures Studies, Stockholm; the Philosophy Department of Uppsala University; and the School of Philosophy, Australian National University, Canberra. For comments on this work, through the different stages of its progress, I want to thank Per Algander, Emil Andersson, Gustaf Arrhenius, Ben Bradley, Tim Campbell, Erik Carlson, Karl Ekendahl, Neil Feit, Molly Gardner, Matthew Hanser, Mats Ingelström, Magnus Jedenheim, Jens Johansson, Kasper Lippert-Rasmussen, Jessica Pepp, Björn Petersson, Olle Risberg, Simon Rosenqvist, Folke Tersman, further participants of the workshop “Harm: The Concept and Its Relevance” (Aug 2016, Uppsala University), of the reading group and workshop on Lippert-Rasmussen's *Born Free and Equal?* (Jan-Mar 2016, Institute for Futures Studies), and of the Higher Seminar at the Philosophy Department, Uppsala University as well as two anonymous referees of this journal.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

## References

- Alexander L (1992) What makes wrongful discrimination wrong biases, preferences, stereotypes, and proxies. *University of Pennsylvania Law Review* 141(1):149–219
- Altman A (2016) Discrimination. In: Zalta EN (ed) *The Stanford encyclopedia of philosophy*. Metaphysics Research Lab, Stanford University
- Arneson RJ (1999) Egalitarianism and responsibility. *J Ethics* 3:225–247. <https://doi.org/10.1023/A:1009874016786>
- Banerjee R (2008) An examination of factors affecting perception of workplace discrimination. *J Labor Res* 29: 380. <https://doi.org/10.1007/s12122-008-9047-0>
- Bradley B (2009) *Well-Being and Death*, OUP Oxford
- Bradley B (2012) Doing away with Harm1. *Philos Phenomenol Res* 85:390–412. <https://doi.org/10.1111/j.1933-1592.2012.00615.x>
- Brake DL (2001) School liability for peer sexual harassment after Davis. In: *Shifting from intent to causation in discrimination law*, vol 12. *Hastings Women's L.J.*, p 5
- Cotter A-MM (2006) *Race matters*. In: *An international legal analysis of race discrimination*. Ashgate Publishing, Ltd
- Dworkin R (1985) *A matter of principle*. Clarendon Press
- Eidelson B (2015) *Discrimination and disrespect*. Oxford University Press
- Ely JH (1980) *Democracy and distrust: a theory of judicial review*. Harvard University Press

- Eyer KR (2011) That's not discrimination: American beliefs and the limits of anti-discrimination law. *Minn L Rev* 96:1275
- Feinberg J (1987) *The moral limits of the criminal law: volume 1: harm to others*. Oxford University Press USA
- Flew A (1993) Three concepts of racism. *International social science review*, vol 68, Winfield, Kan, pp 99–114
- Gardner M (2015) A harm-based solution to the non-identity problem. *Ergo, an Open Access Journal of Philosophy*. <https://doi.org/10.3998/ergo.12405314.0002.017>
- Goldman AH (1979) *Justice and reverse discrimination*. Princeton University Press
- Hanna N (2016) Harm: omission, preemption, freedom. *Philos Phenomenol Res* 93:251–273. <https://doi.org/10.1111/phpr.12244>
- Hanser M (2008) The metaphysics of harm. *Philos Phenomenol Res* 77:421–450. <https://doi.org/10.1111/j.1933-1592.2008.00197.x>
- Harman E (2009) Harming as causing harm, in: *harming future persons, international library of ethics, law, and the new medicine*. Springer, Dordrecht, pp 137–154. [https://doi.org/10.1007/978-1-4020-5697-0\\_7](https://doi.org/10.1007/978-1-4020-5697-0_7)
- Hellman D (2008) When is discrimination wrong? Harvard University Press, Cambridge
- Holtug N (2002) The harm principle. *Ethical Theory and Moral Pract* 5:357–389
- Hook S (2002) Reverse discrimination. In: Cahn SM (ed) *The affirmative action debate*. Routledge
- Kagan S (1991) *The limits of morality*, Oxford ethics series. In: Oxford university press. Oxford, New York
- Kekeles J (1995) The injustice of affirmative action involving preferential treatment. In: Cahn SM (ed) *The affirmative action debate*. Routledge, pp 193–204
- Khaitan T (2015) *A theory of discrimination law*. Oxford University Press
- Klocksiem J (2012) A defense of the counterfactual comparative account of harm. *Am Philos Q* 49:285–300
- Krieger N (2014) Discrimination and health inequities. In: Berkman LF, Kawachi I, Glymour MM (eds) *Social epidemiology*. Oxford university press, pp 63–125
- Krieger N, Sidney S (1996) Racial discrimination and blood pressure: the CARDIA study of young black and white adults. *Am J Public Health* 86:1370–1378
- Lippert-Rasmussen K (2014) Born free and equal?: a philosophical inquiry into the nature of discrimination. In: Oxford university press, Oxford, New York
- Lord Carey of Clifton, 2009. On behalf of Gary McFarlane [WWW Document]. URL <http://thinkinganglicans.org.uk/uploads/On%20behalf%20of%20Gary%20McFarlane.doc>. (accessed 9.15.16)
- Marrast L, Himmelstein DU, Woolhandler S (2016) Racial and ethnic disparities in mental health Care for Children and Young Adults a National Study. *Int J Health Serv* 0020731416662736. <https://doi.org/10.1177/0020731416662736>
- Moreau S (2010) What is discrimination? *Philos Public Aff* 38:143–179
- Morris CW (1984) Existential limits to the rectification of past wrongs. *Am Philos Q* 21:175–182
- Norcross A (2004) Harming in context. *Philos Stud* 123:149–173
- Northcott R (2015) Harm and causation. *Utilitas* 27:147–164. <https://doi.org/10.1017/S0953820814000314>
- Parfit D (1986) *Reasons and persons*. In: Oxford university press. Oxford, New York
- Réaume DG (2002) Harm and fault in discrimination law: the transition from intentional to adverse effect discrimination (SSRN scholarly paper no. ID 285495). In: *Social science research network*, Rochester
- Scanlon T (2008) *Moral dimensions: permissibility, meaning*. Belknap Press of Harvard University Press, Blame
- Shiffrin SV (1999) Wrongful life, procreative responsibility, and the significance of harm. *Legal Theory* 5:117–148. <https://doi.org/10.1017/S1352325299052015>
- Shin PS (2009) The substantive principle of equal treatment. *Legal Theory* 15:149–172. <https://doi.org/10.1017/S1352325209090090>
- Simon S, Moss AJ, O'Brien LT (2017) Pick your perspective: racial group membership and judgments of intent, harm, and discrimination. *Group Process Intergroup Relat* 1368430217735576. <https://doi.org/10.1177/1368430217735576>
- Swim JK, Scott ED, Sechrist GB, Campbell B, Stangor C (2003) The role of intent and harm in judgments of prejudice and discrimination. *J Pers Soc Psychol* 84:944. <https://doi.org/10.1037/0022-3514.84.5.944>
- Thomson JJ (2011) More on the metaphysics of harm. *Philos Phenomenol Res* 82:436–458. <https://doi.org/10.1111/j.1933-1592.2010.00418.x>
- Waldron J (1992) Superseding historic injustice. *Ethics* 103:4–28. <https://doi.org/10.1086/293468>
- Woodward J (1986) The non-identity problem. *Ethics* 96:804–831