Nordic Wittgenstein Review Special Issue 2026 | DOI 10.15845/nwr.si2026.3756
(prepublication for open review)

Minstrel Shows and Murderous Tautologies: Raimond Gaita on Racist Dehumanisation, Love, and the Personal in Ethics

Joel Backström
jbackstr(at)abo.fi

Abstract

Moral understanding is, Lars Hertzberg writes, "absolutely personal" (2022, 105). Alas, this understanding is attacked and obfuscated by what Simone Weil called "the commonest of crimes" in the moral-existential-spiritual realm: the "idolatry" of attributing "a sacred character to the collectivity" (2005, 76). My theme is how this difficulty manifests in Raimond Gaita's ethics. After introducing Gaita's essentially (inter)personal view of the moral understanding revealed in remorse, I argue that his account of racist dehumanisation conflicts with that view. I then outline an alternative perspective on racism as a determination to exclude-and-denigrate-in-the-name-of-collective-belonging and a fantasy of escaping-one's-own-humanity-through-denying-the-other's. My worry is that Gaita's account, rather than analytically exposing racism's moral-existential confusion, unwittingly comes to echo it.

1. Love's light

Raimond Gaita's seminal work in ethics aims to show that "a certain sense of human individuality" is "internal" to understanding what moral problems are, and that a "proper understanding of the character and importance of human individuality will alter our sense of both the moral subject and those to whom he is responsive", these two aspects being "interdependent" (1990, 123; 118). For Gaita, this interdependence is conspicuously revealed in *remorse*, which – in its lucid, truthful form, not its "many and infinitely subtle corruptions" – he characterises as "a recognition of the reality of another through the shock of wronging her, just as grief is the recognition of the reality of another through the shock of losing her" (2000, 33; 2004, 52). Remorse brackets what standard

accounts of ethics focus on: the moral agent's concern to "do the right thing" given their values, norms, and ideals. The moral drama isn't played out between the wrongdoer and their or society's norms – imagining so inevitably results in parody (2004, xxi–xxii) – but between the wrongdoer and *their victim*. Remorse opens us to "a sense of the significance of the evil we did and a sense of the reality of our victim", one through the other (2004, 147).

Remorse, then, is "the pained recognition of [...] what it means to have wronged someone", where "the meaning of what one has done, what one has become through doing it, and what one's victims have suffered, are inseparable" (2000, 4; 34). Furthermore, this recognition is possible only insofar as one sees one's victim in the light of love, for "something is precious only in the light of someone's love for it" (2004, xxiv). When one looks at another lovelessly, as hateful or indifferent, one's sense of their preciousness, and so of one's actions as *violating* them, is pushed aside. Indeed, love is, Gaita says, the "most fundamental concept" of morality, insofar as we wouldn't "have a sense of the sacredness of individuals, or of their inalienable rights or dignity", were it not for "the many ways human beings […] love one another" (2000, 8; 5; cf. 2004, 146; 2000, 26).

Gaita's own work is "inspired" by a nun he met while working in a psychiatric hospital as a teenager (2000, 2). Whereas everyone else treated the incurable patients brutishly or, at best, with benevolent condescension, she responded to them "with a love of such purity" that it managed "to reveal the full humanity of those whose affliction had made their humanity invisible" (2000, 20). Gaita writes "as someone who was witness to the nun's love", and he tries to articulate the fullness of relationship between individual human beings it revealed; "to mark its conceptual features, to locate it in a sympathetic conceptual space" (2000, 21; 2004, xxxi).

Lucid remorse is impersonal insofar as it reveals that mistreating someone who is, socially, a "nobody", or whom one personally dislikes or despises, is as terrible as mistreating the most "important" or "worthy" person; in that sense, one's victim "could have been anyone" (cf. Gaita 2004, 147–50; 2000, 30–2). But precisely because this is so, one can also speak of "the individuating work of remorse which ensures […] that our victims remain with us in all their individuality" (2000, 32) and, as the other side of this, of "the radically singular 'I' that is discovered in remorse" (2004, 55). In my remorse, I'm lacerated by the realisation that it was *I* who did this to *you*; I see your face before me, I hear

your cry from the pain I caused or remember the hope in your eyes that I later betrayed. As I recall these or similar things in the light of love, I'm alive to you in your individual reality, rather than seeing you in terms of my own and my community's preferences and valuations concerning who is and isn't important, and how. And that's also how I find myself as a human being. That is, as I would say, now going beyond what Gaita explicitly says, but drawing out what to me is its central implication: I find my soul, open to the other's soul regardless of the social persona and psychological personality defined through those valuations and preferences, through that whole "game" of sympathy and antipathy, of alliances and aversions that structures collective life. In the shock of remorse, I realise that I had given myself over to that game's loveless logic of like and dislike, alike and unlike, of "me" vs. "others" and "us" vs. "them", thus deserting my sense both of myself and of you (cf. Backström & Nykänen 2016).

2. Racism: the light turned off

Gaita's conception of remorseful interpersonal moral understanding contradicts standard philosophical conceptions of ethics which, he notes, speak of principles, values – or of whatever, really, so long as *the other person*, the victim of one's actions, "drops out and becomes merely an instance of something else that carries the moral weight" (2004, 147–8). However, Gaita gives no indication that here, philosophical theories echo the "idolatry of collectivity" (Weil 2005, 76) in everyday life, whereby collectively elaborated moral norms and sensibilities are taken to determine the moral significance of what transpires between oneself and individual others. Typically, this *depersonalisation* (Nykänen 2009) doesn't take *obviously* impersonal or conventional forms, nor is it experienced as submission to an external collective authority. Indeed, the more one identifies as 'one of us', the less will one consciously feel any conflict with one's community, as one will "say [...] *I* with a collective signification" (Weil 2001, 12).

Gaita doesn't thematise this problematic, but rather inclines – despite his account of remorse and despite Simone Weil being the thinker he quotes "more than anyone else" (Gaita 2014, 177) – towards what appears to me an essentially collective view of moral meaning. This becomes evident in his discussion of racism, centrally placed in his work and intended "to illustrate the role that the concept of a human being should play" in our efforts to

understand moral concepts (2004, 331). In my view, this amounts to a crucial, unresolved tension at the heart of Gaita's ethical thought between tendencies towards strictly interpersonal and towards collectivist articulations; a tension that Gaita himself apparently doesn't feel, however.[1]

The basic contention in Gaita's discussion of racism is that the power of others to affect us morally is "not primitive" but presupposes that we "take them seriously", that we "find it intelligible to ascribe certain kinds of thoughts and feeling to them" – as we may not always do (2004, 156). He claims, for example, that racist white slaveowners of the American South found it "unintelligible" that certain moral descriptions should apply to their black slaves; those that "mark our sense of what it is to be a fully human being" (2004, 153; 157). Slaveowners couldn't see their slaves as "an intelligible object for anyone's love" (2004, 161), and so couldn't see their own actions towards slaves as *violations* – which is why, or how, they could enslave them. To be "within the conceptual reach" of our moral responses, human beings must be seen as "*'one of us'*, a fellow in a realm of meanings which condition the way we may matter to one another", and since slaveowners didn't see slaves that way, whatever they did to them was not "within the intelligible reach of [their] remorse" (2004, 157; 151, emphasis added). For example, a slaveowner could rape a black slave girl *without feeling a thing* in the morally crucial sense, and since he'd feel no pangs of remorse, there'd be nothing for him to hide from himself through self-deception, either – as there would be, Gaita says, if he raped a white girl (2004, 160).

Gaita claims, then, that for the racist, individual human beings may cease to exist, morally speaking, their humanity rendered "epistemically impotent" (2004, 162). He underlines that racists generally don't "suffer from ignorance of what we ordinarily call facts about the victims of their denigration", and slaveowners certainly treated slaves quite differently from animals (2005, 86; cf. 2004, 156–8). They didn't prohibit their horses from teaching themselves to read, for example (cf. Douglass 2009, 44–5). Nonetheless, Gaita insists that racists don't see their victims as fully human. In raping a slave girl, the

[1] Gaita discusses racism in both *Good and Evil* and *A Common Humanity* (see especially 2000, 57–72; 2004, 156–163; 331–41). He warns against equating "all forms of racism" – he thinks 'colour-racism' differs essentially from antisemitism, and both differ from "ethnic hatreds" – and acknowledges that his account omits the "many psychological needs which racism satisfies" and so doesn't constitute an "adequate psychology of racism" (2000, 68–72). Nonetheless, he wants to show "how rich and […] radical the lessons of an analysis (conceptual) of a certain kind of racism can be for moral philosophy" (2004, 331).

slaveowner need not, Gaita holds, be blind to the suffering in her eyes, and may even feel ashamed for his lack of pity – but then he may have felt similarly ashamed for his cruelty "when he beat his dog" (2004, 188). The point, for Gaita, is that the slaveowner cannot see raping a slave as a violation that would call forth remorse rather than mere shame (cf. 2004, 339). And insofar as our "sense of the reality of another" is "conditioned" by "the individuating work of remorse", to lose the very possibility of remorse in relation to others is "to lose a sense of [their] full humanity" (2000, 34; 32; 5).

Gaita apparently considers this account of racism fully compatible with his claim that remorse, the loving response of one person to another, "discloses the fundamental determinant of our understanding of what it is to be a human being" (2004, 151). He simply adds – but everything hangs on this 'simply' – that the very possibility of remorse arises only once particular individuals have been put "within its conceptual reach" by "being seen to be 'one of us'" (2004, 151). In Gaita's view, remorse is nonetheless still morally "fundamental", insofar as its nature is "underdetermined" by, and indeed "radically transforms" what – namely 'our' collective sense of shared humanity, from which certain others may be excluded – "conditions it"; for example, one's sense of what it means to be a friend or husband may be "transformed under the shock of what a human being is disclosed to be in serious remorse" (2004, 151).

Gaita suggests, then, that whether human beings 'register' to one's moral perception at all is decided by one's collective identification, that is, by the way one understands oneself to belong to a particular community that excludes certain 'kinds' of others as 'not one of *us*.' Once – *if* – someone is allowed to 'show up on one's moral radar', however, *then* that person is invested with extraordinary powers to transform one's understanding of the collectively elaborated world of meaning that decided whether one would 'see' her at all. Correspondingly, racists may display the deepest and clearest moral sensitivity and understanding imaginable vis-à-vis those they see as 'one of us' – yet may lack any sense of moral connection to their racialised victims.

I cannot make sense of this suggestion. And as far as I can see, Gaita offers no reason why we should – or philosophical articulation to show that we coherently could – think of racism and moral understanding in this bizarrely split way. He simply insists that we should. Real-life slaveowners have certainly often *said* that slaves aren't fully human, as racists more generally have said this

of their victims, and history abounds with highly cultured and otherwise sensitive slaveowners and racists. But that doesn't settle what saying such things *means*, how one is really relating – what one is doing – to others and oneself in saying or thinking such things. These are the philosophical questions I want to discuss. I suggest we try to understand – not condone – the actions of racists on the assumption that, like evildoers generally, they know (feel, understand) that their victims are fully human and that they are wronging them, but wrong them nonetheless, from familiar motives such as greed, fear, vengeance – and, crucially, from a felt need for *collective belonging*, for being 'one of us' even when that entails wronging 'outsiders.' Like people generally, racists deny or excuse their wrongdoing not because they find remorse 'unintelligible', but precisely to *defend* against the remorse, the bad conscience, occasioned in them by victimising others. It isn't that they don't know they do evil, but that admitting it is unbearable (cf. Backström 2023). I will indicate how these familiar dynamics can be seen at work in racism. Like Gaita, I focus on racists rather than on the experience of their victims; like Toni Morrison, I'm interested in "the impact of racism on those who perpetuate it" (1993, 11). This implies no disregards for the victims, and that impact is itself the effect on the souls and worlds of racists of what they do to their victims.[2]

3. The minstrel show

The white racist, Gaita claims, sees black faces as if 'in blackface', that is, as inherently caricatures of the human form. Wittgenstein said that the human body is the best picture of the human soul, and faces seen as "like the *Black and White Minstrel Show's* caricature" cannot, Gaita notes, "picture souls that have any depth" (2014, 174–5). "The racist's thought", Gaita says, is that "that is how they look, *essentially*", and that black people look like that to him is "fundamental to what makes them 'them' and to why he finds it inconceivable that they should be treated as 'one of us'" (2000, 61). There's no need to invoke

[2] Space prevents discussion of the complex debates, from Fanon 2008 to Hartman 2022, over how racism, colonialism, slavery and their aftermath affect their victims; the suffering, resistance and witnessing they provoke and constrain. I also cannot discuss the problematic, actualised by Gaita's use of rape as an example, of gendered and sexualised violence in general, and specifically in the context of American racialised slavery (cf. Hartman 2022, hooks 2015). Finally, while I predominantly focus on (North-)American slavery because of its centrality to Gaita's discussion, racism and slavery exist in countless variations, and focusing on other examples would likely bring out interestingly different aspects of the general problematic.

stereotypical (empirical) beliefs or bigoted 'moral' principles to ensure that nothing serious can come from faces seen that way. *That perception itself* places the inner lives of black people "in inverted commas", as Gaita says: "We grieve, but they 'grieve' […] we love and they 'love', we feel remorse, they feel 'remorse' and so on" (2000, 63). Such faces and such 'inverted commas' expressions are, in Gaita's view, *all* racists ultimately perceive in the racialised other; he explicitly denies that racists "know in their hearts that their victims are fully human" (2000, 72).

To Gaita, the "most important point" (2014, 175) of his analysis of racism is the claim that racists see their victims this way; that is, that they suffer from a "meaning-blindness" that makes them unable to see in their victims "the possibility of *ever deepening* responses to the *meaning* of facts of the human condition" (2005, 86; cf. 2004, 333). Gaita finds reflection on racism thus conceived philosophically instructive because he thinks "moral philosophy is often characterised by [the same] kind of meaning-blindness" (1994, 626). Racists as depicted by Gaita attribute to their victims "all the raw materials from which philosophers have […] constructed theories of morality" – for example, "that they are rational, have interests, that indeed, they are persons" – but since racists nonetheless fail to see their victims' (full) humanity, "their example puts that philosophical tradition seriously to the question" (2004, 339).

I agree that moral understanding, and our difficulties with it, are most basically manifest in 'full bodied', meaning-laden responses involving one's whole person in one's relation to others. Gaita seems to forget, however, that blackface minstrel faces are products of active *caricaturing* by racists bent on denigrating black people rather than truthful (re)presentations of how they saw black people. Had white audiences actually seen black people in the way minstrelsy caricatured them, there would have been no need for the caricature; they could then simply have watched black people dance, sing, and make jokes in the way they did when among themselves. Instead, they needed the denigrating caricature, and when actual black people – as opposed to white performers in blackface – were allowed onto white stages, they had in effect to

take on the minstrel mask, too, turning themselves into caricatures; be buffoons, happy simpletons.[3]

Gaita is right that depth, seriousness, full humanity, is missing from the minstrel face; the problem is that he presents as simple *absence* what racists had studiously *removed* from the black face. It isn't that whites couldn't see those aspects in black faces and bodies, but that they didn't *allow* them to appear. Blackface minstrelsy was an ideological ritual, and such rituals are primarily "a kind of self-hypnosis within ruling groups" (Scott 1990, 67). Minstrel characters were, as Ralph Ellison said of black characters in white American fiction generally, "key figure[s] in a magic rite by which the white American [sought] to resolve the dilemma arising between […] his acceptance of the sacred democratic belief that all men are created equal and his treatment of every tenth man as though he were not" (2003, 85). In other words, the inverted commas white racists place, as Gaita says, around the 'inner lives' of black people are really projections of the inverted commas they have placed around their own 'perception' of them, because admitting the full humanity *they themselves actually see* in black people would force them to see what they have turned *themselves* into in treating black people as they do.

My claim isn't merely that a need for denigration in the service of racist oppression was the central socio-psychological *motive* or *cause* of the representation of black people in minstrelsy; a point Gaita himself concedes (2004, 334). Gaita distinguishes between the "psychology of racism", concerned with "the many psychological needs which racism satisfies" – with which Gaita says he *isn't* concerned – and his own project of describing "the conceptual structure […] of a certain kind of racist perception" (2000, 72; 2004, 335). My point, however, is that this distinction collapses, because the racist's very perception of the other is *structured*, not just caused or motivated, by racist 'psychology', that is, by the will-to-denigrate manifest in racist ideology. The minstrel face is *essentially* a denigrating caricature: (something like) that is what human faces are turned into when the will-to-denigrate deforms one's perception. No-one can – they conceptually cannot – look like that 'in themselves'. What the minstrel face shows is really the gaze directed onto it. In other words, *that face pictures the racist, not his black victim*. As James Baldwin said

[3] Blackface minstrelsy, most popular from the 1830s to 1870s, was no mere curiosity but the first truly national American popular-cultural entertainment form, hugely influential on later popular culture and music, white *and* black. There's a large literature on its multifaceted, contradictory, and contested history and influence; see, e.g., Bean, Hatch & McNamara 1996, Lhamon 2003, Lott 2013.

of racist denigration generally, "You have not described me when you call me a nigger […] You have only described yourself" (2010, 61).

Denigrating others is essentially projection, and essentially ambivalent. The very urge to denigrate reveals a fascination with the other, denigration being a mode of engaging with them, however destructively. As historians of blackface minstrelsy agree (cf. references in footnote 3), its representation of black people wasn't simply and wholly negative, but also expressed white identification with and envy of black people, manifest in the very enjoyment of imitation, of 'taking on' characteristically 'black' moves in song, dance, and speech. Minstrel entertainers "claimed to be pupils, or even kin, of the Blacks they mocked" and "occasionally addressed [their white audiences] as 'my broder [brother] niggars'" (Roediger 2007, 116–7). Crucially, however, these (more or less) positive aspects were overlaid and undermined, limited and perverted, by the dominant ideological need to denigrate present in minstrelsy; specifically, by "the fierce insistence by the white ruling class that Negroes do nothing which might lead either themselves or other people to believe that they are equal or superior to whites" (Cox 1959, 366–7). Like interracial interactions in racist society generally, minstrelsy was allowed to unfold only where "the relation of master and servant [was] visibly established between race and race"; thus, white people could lampoon black people, not the other way around, and the 'black' man on stage could only make fun of himself, never of white people; *then* there could be white 'fun', "clouded only by […] the perpetual unrest that always accompanies forcible possession of anything" (Cable 1903, 21–3).

4. A murderous tautology

Gaita claims that white slaveowners found it "literally unintelligible" that black slaves could be wronged, that they genuinely were not "*able to hear*" and "*could* not have heard" a fully human address in black voices or seen it in black faces (1994, 623; 2004, 160; cf. 2004, 333). Supposedly, this mysterious inability – mysterious because left quite unexplained by Gaita – explains how they could enslave black people.[4] I would turn this around: *because* they enslaved black people, white slaveowners couldn't allow themselves to *acknowledge* their full humanity but tried to obfuscate it through the elaboration of a mythology and choreography of black inferiority; minstrel shows of all kinds, whereby a

[4] Gaita says that *if* a slaveowner "could be haunted" by the evil he did to his slave – as Gaita thinks slaveowners generally could *not* be – "then her days as a slave would be numbered" (2004, 157).

fictitious 'essential difference' between the 'races' was embodied in social reality. Or will anyone seriously suggest that Southern whites just happened to have a collective predisposition that made them unable to see black people as fully human, and that this 'happy accident' (from the point of view of slaveholding) made it easy for them to *then* enslave them?

Certainly, racial and ethnic prejudice can be easily provoked between human groups, us/them-thinking and hostility towards the foreign(er) being ever-present human potentialities – just as are sympathy with and interest in others. But ephemeral prejudices within changing constellations of group-contact and conflict should be distinguished from racism proper: a settled, institutionalised hostile sense of fundamental difference between 'us' and a particular racialised 'them' of the kind Gaita describes. In the case of Gaita's main example, (North-)American slavery, as a matter of historical fact racism didn't predate black slavery but was developed as its ideological justification (cf. Fields & Fields 2012, 121–8; Fryer 2018, 135–6). Enslavement, like other evils, needs justification, and if slaves but not their masters happen to be black, denigrating them based on their skin-colour conveniently provides one. To buy and sell white people like cattle would be quite unacceptable, white slaveowners told themselves, but since, in the words of one 1859 apology, "nature itself has assigned his condition of servitude to the Negro [...] he is not robbed of any right, if he is compelled to labor" (quoted in Du Bois 1964, 52). Alberto Memmi summarises this basic racist logic: "If I dominate you, it is *because* you are an inferior being; the responsibility is yours, and the differences that exist between us prove it"; in other words, racism means "charging the oppressed for the crimes […] of the oppressor" (2000, 56; 139).

Obviously, racism exists in contexts much less oppressive than slavery, just as slavery and other forms of systematic oppression can exist without racism and may be 'justified' by invoking, say, 'essential' differences between the sexes or between religions instead. Whatever the variations, however, a hostile determination to exclude and denigrate those branded Others is part of what we mean by calling out something as 'racism'. And whatever further violence and oppression racism issues in – the whole machinery of guns and police, of 'whites only' signs and yellow stars – this basic hostility itself needs justification to appease the racists' own unease over the hostility they direct against their victims. As can be seen in example after example, and as I try to illustrate here, there's no oppression of others without repression of the oppressor's own

conscience (cf. Backström 2019; 2023). Mythologies presenting 'them' as inferior and to-blame for their own oppression are the vehicle of this repression, even as they are also an essential instrument and expression of the very hostility whose evil they obfuscate.

I thus disagree with Gaita's claim that, although contemporary racism is "on the defensive [...] in the face of widespread anti-racist sentiment", the racism of American slaveowners, for example, was non-defensive (2000, 65; cf. 68). Racism is *essentially* defensive *because* it's essentially aggressive, and racists need to defend both against their victims and against their own unease over victimising them. This aggressive defensiveness is manifest in countless ways, for example in racists' responses to those who treat their victims with respect. Consider Gaita's description of the staff at the psychiatric hospital where he met the nun. Although no 'racial' dimension was involved, Gaita says the staff "found it as unintelligible that such [incurable] patients should be accorded full respect" as white racists find it unintelligible to respect black people, and they responded with "fierce scorn" to the few progressive doctors who endeavoured to treat the patients respectfully; a scorn "expressed in the same tone that the accusation 'Nigger Lover'" had been thrown at whites who treated black people respectfully (2011, 56). Gaita reports this, but fails to see that it reveals how, contrary to his central claim about racism, the staff didn't, any more than racists do, find it *unintelligible* but rather *intolerable* that their victims should be treated respectfully. The response to what one finds genuinely unintelligible is bafflement, simple incomprehension, whereas one responds to what one finds intolerable – precisely *because* one *understands* what it implies – as the staff did, with violent resistance, ranging from ridicule and scorn all the way to murder.

There are all kinds of criss-crossing (statistical) differences between human groups, however defined, as between individuals within them; this is just a fact about human variability. But the essential (capital D) Difference that racists claim exists between 'us' and 'them' ultimately has no other content or ground than the hostile determination to *make* it, to separate 'us' from 'them'. And Gaita to the contrary, there's nothing particularly special about "racism directed against people whose skin colour and whose facial features are significantly different" (2000, 65; cf. 72). Even where such objective physiological differences exist (more or less) between oppressors and oppressed, the Difference that racists make such differences signify – which

supposedly makes it 'impossible' for 'us' to understand 'them' or makes it 'inevitable' that 'we' should rule over 'them' – exists, again, only in and as the determination to *make* it. Conversely, even where oppressors and oppressed are of the same ethnicity, oppressors tend to project an essential Difference between themselves and their victims. Thus, Russian noblemen presented their equally Russian serfs as "intrinsically lazy, childlike, and requiring of direction", and sometimes even "claimed that whereas they had white bones peasants had black bones" (Kolchin 1987, 170–1). And George Orwell, speaking of (practically) all-white English class-society, summed up "the real secret of class-distinctions in the West" in the middle-class conviction that "*The lower classes smell*": "The smell of their sweat, the very texture of their skins, were mysteriously different from yours" (1962, 112–3).

In other words, the concept 'race' that figures in racism is purely mythological. It isn't just that races are 'socially constructed' but that 'races' are *created by racism*. Racism is a reality, a real crime; 'races' are an illusion, a lie created by, and as the vehicle of, the crime (cf. Fields & Fields 2012, 101). In effect, Gaita acknowledges this when he notes that the racist can give no determinate characterisation of the supposed Difference between 'us' and 'them': ""They' can do and feel almost everything we can *except not as we do*" (2000, 63, emphasis added). The racist cannot explain "why 'they' are not human beings, but he most definitely means that 'they are not one of us'" (2004, 159). That is, indeed, the empty core of racism, its murderous tautology: "The difference between 'us' and 'them' is that they are not us".[5] In my view, Gaita's *confusion* lies in thinking that the racists' insistence on this 'essential difference' is itself grounded in, or simply consists in, a perception of their victims supposedly characterisable without reference to the determination-to-exclude-and-denigrate and the dynamics of defensive projection indicated above. That isn't a philosophical analysis of racism but – although Gaita certainly doesn't intend it that way – a restatement in philosophical language of the fundamental racist myth, of its murderous tautology.

[5] The racist community is as hollow as is the concept 'race.' Racists are united only by their shared hostility towards racialised 'others', and if these others disappeared, the racist community would dissolve (cf. Sartre 1995, 20; 8).

5. The ring of Gyges and the love of saints

The racist's mythological representation of his victims as not fully human allows him to pretend that they deserve, or don't suffer from, their treatment, but also that they cannot see and (fully, really) understand *him*. American slaveowners thus staged a "pretense of invisibility" through "white control of the black gaze", with slaves brutally punished for looking at or appearing to observe their masters (hooks 2015, 168). Whereas slaves had no illusions about their oppressors, the latter needed to deny their victims' negative, disdainful responses to *them, their masters*, and thus prohibited overt expression of such responses (cf. Blassingame 1979, 256–7). Allowing oneself to realise that one's victims understand the evil one does and see through one's 'justifications' would make one feel, in Hannes Nykänen's phrase, "unbearably naked", and insofar as one is determined to do evil one must therefore repress *one's own knowledge that the other knows* (2019, 350). Thus, every evildoer dreams of the ring of Gyges, not primarily to hide his crimes from 'society', but to hide himself from his victims and so from himself.

Consider a racist bully attacking his victim. The bully feels safe in and is egged on by his power over the terrified victim. But now suppose the victim – perhaps a slave about to be raped by her owner – looks the bully straight in the eye, showing him that whatever he's about to do to her, and although he's physically stronger and has slave-law on his side, so that in worldly terms she's powerless to stop or hurt him, *she's not afraid of him* and sees through the evil futility of his power-games. This takes the ego-boosting excitement and sense of impunity out of the bully's violence. He may react in different ways. Her look, her response to him, may awaken him to what he's doing; he may remorsefully seek her forgiveness and be led to change his life. Or he might merely feel ashamed, leave her alone and perhaps sell her down the river to avoid being reminded of this personal humiliation – which is how he misrepresents the encounter to himself. Or he may turn even more violent precisely due to his desperate realisation that violence will get him nowhere with this human being who, by just being there, unafraid, reveals his own futility more clearly with every blow he deals her. And so perhaps he ultimately

kills her; kills her rather than *facing her and facing himself*, what he has made of himself.[6]

The destructiveness exemplified by that last response reveals the anguish engendered in our struggles with conscience, with opening oneself to another human being. Gaita claims – without showing why we should, or how we could intelligibly think – that the possibility and difficulty of this encounter could somehow be eliminated by racist collective identifications and dehumanisation. I propose, instead, that we view racism's existential 'tendency' in the light of the core fantasy of making one's own humanity – along with what one does to it in dehumanising others – invisible, to oneself and to them. Philosophers perpetuate this fantasy by reducing human encounters to what cultural conceptual schemes officially allow, claiming that "only by virtue of certain kinds of [...] cultural frames will a given face seem to be a human face to any one of us" (Butler 2005, 29–30). This reduction subordinates (inter)personal moral understanding to the rule of collectivity, thus eliminating – in conceptual fantasy – the conflictual tension between them.

This reductionism threatens in Gaita's discussion of love, too. Despite the apparently radical claims quoted in Section One, Gaita ultimately presents love quite conventionally, as a family of natural passions and attachments, some of which are destructive and "in conflict with one another and with morality" – which means that morality is "in tension with what [i.e., love] conditions its most fundamental concept [i.e., the infinite preciousness of human beings]" (2000, 8; cf. 25–7). To play the fundamental role that Gaita assigns to it, love must therefore be of a particular kind, judged in the light of "standards" to which "lovers must try to rise"; standards that, while "internal to […] love itself" in the sense of determining our concept of love's "real as opposed to its counterfeit forms", are nonetheless cultural creations (2000, 24–5). Specifically, Gaita says that love of the kind the nun showed the patients "would not be possible if her culture did not make it possible for her"; in the absence of the tradition or "language" of "saintly love" we would "not have had an idea of what the nun's love revealed and her love *would not have the same thing to reveal*" (2004, 123; 2011, 64, emphasis added). Gaita suggests, then, that *humanity* in

[6] Standing up to bullies may profoundly change oneself, not just the bully. Frederick Douglass describes his decision to finally fight back against his slave-master as "a glorious resurrection"; "I now resolved that, however long I might remain a slave in form, the day had passed forever when I could be a slave in fact. I […] let it be known […] that the white man who expected to succeed in whipping, must also succeed in killing me" (2009, 78).

the sense most crucial to moral understanding – the irreplaceable individuality and preciousness of every human being – is itself a product of, was "bequeathed to us" by, its revelation through a contingent process of cultural elaboration (2018, 217; cf. 2004, xxii–xxiii).

Now, what's wrong with this? Certainly, there's a "language of love" which "nourishes and is nourished by our sense that human beings are irreplaceable" (Gaita 2004, xxiii–xxiv). Insofar as human beings both speak and love, we express our love also, and centrally, in words – as in glances and caresses, in actions and responses of all kinds, all of them addressed to and oriented by those we love, love being itself all expression, all reaching-out-and-opening-oneself-to-the-other (cf. Backström 2007, Nykänen 2002). But I see no point in asking, nor any way to determine, how much of love's expressive movement could exist 'without words', or in the absence of this or that particular tradition of cultural articulation, say that of Christian ethics and saintly love. The salient point is that, regardless of which words are used or traditions drawn on, no love will be expressed unless one actually *loves* those to whom one turns. Think of the nun. Not 'the tradition of saintly love', but *her love* – that is, the way she turned *to the patients*, "everything in her demeanour towards them" – showed them in a new light to Gaita, and revealed to him that, while he and others spoke of the patients' dignity, "in our hearts we did not believe this" (2000, 18–19). And if the nun was as Gaita depicts her, *she* would certainly never say that people outside of her religious tradition could not see and love others as she could.[7]

6. Sentimental depths

In Gaita's view, racists genuinely believe their own claims about their victims' deficient humanity; they "mean what they say", and their apparent inability to see the evil in their own actions is "genuine" (2000, 63). But how does Gaita know that? And, more importantly, what is the sense of this kind of claim? If the contrast to 'genuine' belief is the simple lie or pretence, racists no doubt often genuinely believe what they say, but that contrast is far too crude for describing the workings of racism, or moral-existential difficulties generally, where self-deception, repression, and self-suggestion are in play.

[7] This isn't to deny that, as Gaita stresses (cf. 2000, 17–27), various discourses in and out of philosophy make it virtually impossible to say anything meaningful about love in their terms.

Consider this declaration from a real-life American slaveowner (not Gaita's imaginary character), who says that "there are few ties more heartfelt, or of more benignant influence, than those which mutually bind the master and the slave" – but who, a mere page later in his apology, adds a stern warning that if abolitionist propaganda were allowed to spread, and these ties were loosened even for a day, the slaves would immediately proceed to massacre their 'beloved' masters (quoted in Elkins 1976, 219, and fn 136). This is absurd, but in another sense expresses the moral reality of the slaveowner's situation quite exactly. Like moral criminals generally, he needs to pretend that what he's doing to others is *not* a crime but something good. But although he officially denies it, he knows that he is indeed committing crimes and that his victims resent him for it. Thus, he's deadly afraid of what would happen were the social order of slavery to break down, and he expresses that fear frankly. Since this isn't an expression of moral insight but a self-deceptive apology for crime, however, he does so in the disguised form of a schizophrenic portrait of his victims, split into the loving, childlike 'good slave' and the rapaciously violent 'bad slave', with himself presented as an all-good figure who must, sadly, also arm himself to the teeth because the monstrous 'bad slave' is unaccountably present within the 'good slave'. This is all said quite clearly by the slaveowner, and yet 'unconsciously', that is, without him being willing to admit the true meaning of what he's saying, the terrible moral reality he's revealing. If he admitted it, his apology would turn into self-accusation, would be revealed to him, too, as what it already clearly is to us, his readers.

Did the slaveowner mean what he said? Well, he must have convinced himself that his apology – a far from original one – would be taken seriously.[8] One can get oneself to 'believe' all sorts of nonsense, but one cannot lucidly *mean* nonsense, however convinced one may feel of its sense and truth. One feels convinced because one *needs* to feel convinced, because one is deflecting and denying responsibility for wrongdoing, repressing the unbearable bad conscience that doing evil to others occasions. The slaveowner repressed his conscience by working up feelings of resentful self-pity against the 'ungrateful' slaves threatening his sentimental family-idyll. There's nothing genuine about such feelings; they result from self-suggestion – although, for that very reason, they aren't insincere either, as one stages the same show for oneself as for

[8] In American proslavery writings, "a paternalistic insistence on the humanity and harmony of slavery was as pervasive as the racial argument in its defense" (Kolchin 1987, 172).

others. To give oneself over to sentimentality is to falsify, and it feels comforting precisely because one renounces all attempts to confront the reality of one's relations to others. One cannot be *genuine* in being sentimental, and the more deeply sentimental one allows oneself to be, the falser one becomes.

Gaita often refers to sentimentality in discussing the concept of falsity applicable to questions of (moral) meaning. Moral reflection, he says, has as "intrinsic to its content, judgements that this or that is [e.g.] sentimental, or [...] banal" (2004, xxxiii). I don't think he ever mentions sentimentality in connection with racism, however, although a deeply sentimental orientation pervades much racism.[9] Sentimentality is essentially a misrepresentation of love, simultaneously prettified and cynical because truth and justice are refused, and love is thus replaced by mere affection or adoration, and by an oppressive 'belonging' that – as the slaveowner's 'idyll' illustrates – doesn't challenge but instead demands allegiance to the injustice and violence of one's community.

As Gaita notes, where sentimentality is seen as "a form of the false", so that a response being sentimental is "what is primarily wrong with it", no "discursively establishable principles" allow us to determine what is and isn't sentimental (or cruel, self-pitying, etc.), but we must simply judge in the light of "what has moved us in the speech and actions of others" (2005, 51; 2004, 269–70). This means that one must ultimately speak "in the first-person singular", even as what one says concerns others as much as oneself (2004, 198; cf. 205). It also means, insofar as we credit Gaita's claims about the centrality of love for moral understanding, that one can only "bear witness" to what one sees revealed as good and evil in the light of love (2014, 181). If someone claims not to see or to disagree with what one sees – for instance, seeing no falsity or sentimentality in the slaveowner's apology – one cannot *prove* them wrong, yet that doesn't make the moral meaning or truth of what one sees uncertain. And if you now ask why not, I'd ask, not in the spirit of proving anything, but as a real question: "Well, *is* it uncertain for you?" I'd also remind you that anyone who wishes to defend oppression *must* claim to find nothing wrong with it, and that there's no position from which to view oppression while remaining 'uncommitted' as to whether it *is* oppression, for that supposed 'neutrality' means siding with the oppressor against the oppressed, passively looking on while the oppression continues (cf. Backström 2023, 90–2).

[9] Kipling's poem "The White Man's Burden" is a case in point.

7. "Dey wasn't a soul in de whole place"

What I find confused in Gaita's perspective on racism also leads him, it seems, to misdescribe what freeing oneself from racism may mean. He says that, for his generation of (Australian) white people, raised watching films portraying black people as brutish and ridiculous, the sixties brought "a revolutionary change in sensibility" as they realised that black people were, as Gaita often says in related contexts, "fully our equals": "Like us they love rather than just 'love' […] Like us they are unique and irreplaceable" (2014, 177; 2000, 3). But if we wish to speak of equality, shouldn't we rather say that, in giving up his racism, the ex-racist finally shows himself the equal of those he denigrated? For there was never anything amiss with *their* humanity; by denigrating them, *he* made himself inhuman.

This reversal of perspective also leads one to reconsider the character of the "realm of meaning" from which Gaita's racists exclude their victims; the realm in which "we discuss what it means to love or grieve truthfully and why it matters; what it means to suffer wrong and what it means to do it", etc. (2014, 176), and in which "our sense of […] good and evil and […] of what it means to live a human life" can deepen in "a dialogical engagement" between us (2004, 341; 338). Gaita writes as though this "conversational space" remained unaffected by the exclusion of some people, denigrated as not fully human, from it (2004, 338). But this isn't so. Consider, for example, Gaita's claim that slaveowners thought of themselves as Christians, but of their slaves only as 'Christians'; that is, that they thought only of themselves "as serious respondents to the question, 'What is it to be a Christian?'" (2004, 159–60). Slaveowners indeed often presented matters thus, but Gaita fails to address the moral confusion and irony in that thought. For it's clear – and it was always clear to the *slaves* – where the inverted commas belong when 'Christians' are inside their Churches praising neighbourly love on Sundays, while every day they're enslaving their fellow human beings. While slaveowners constructed arguments for the 'Christian' nature of slavery (Fox-Genovese & Genovese 2005, 505–27), slaves thus told different stories. In one, the master, dismayed, tells his slave he dreamed he "went to Nigger Heaven" and saw "the muddiest, sloppiest streets […] and a big bunch of ragged, dirty niggers walkin' around". That's curious, his slave Ike answers, "I dreamed I went up to de white man's paradise, an' de streets was all gol' an' silver […] but dey wasn't a soul in de whole place" (Lomax 1960, 464; cf. Bay 2000, 178–83).

In barring black people from white churches, then, white people weren't depriving them of Christianity but excluding Christianity from their own lives, thus emptying their cultural 'realm of meaning' of meaning. This illustrates how racism cannot be confined to the racists' direct dealings with their victims: its corrupting influence spreads – at once pervasively and unpredictably in its ramifications – into the racists' life generally. As Ralph Ellison said: not despite but *because* of racial segregation, "Southern whites cannot walk, talk, sing, conceive of laws or justice, think of sex, love, the family or freedom without responding to the presence of Negroes" (2003, 163; cf. Morrison 1993). Racism cannot be contained, 'kept in its place' like the denigrated and oppressed it is designed to keep in theirs, because violent oppression exercised over *some* people must be extended to anyone who, and anything that, threatens or challenges that oppression. The deeper point is that racism is a form – one among many – of collectivist depersonalisation; in other words, a refusal and repression of love, an attempt to limit love, to keep *it* in its 'place', as defined by various collectively sanctioned norms. But love, the concern and longing between human beings, has no proper place; it's the very movement of the human soul that takes us out of any place we may consider 'ours' by opening us and setting us moving towards each other – towards *each* other. When we try to contain this movement, to exclude from it certain 'kinds' of people, their exclusion ultimately remains an illusion, kept up in the form, and at the cost, of the never-ending, literally soul-destroying work of hardening and crippling ourselves.[10]

Thus, a child who becomes aware of his parents' active participation in an oppressive racist culture – like 8-year old Jesse, taken by his parents to a lynching in Baldwin's (1965) short story 'Going to Meet the Man' – faces a terrible choice: *either* recognise that your parents have given themselves over to evil *or* convince yourself that joining the oppression is part of what 'love' *means* for you and for them. This 'love' is counterfeit, with openness to others replaced by a conspiratorial intimacy violently excluding some (here, black people) and prohibiting any questioning of that exclusion by those 'within'. Jesse's parents thus withhold love's openness from their son as much as from black people, even as they *present* what they do as love for him and for their community (cf. Nykänen 2002, 207–17). The child knows in his heart that this

[10] The analyses in Nykänen 2002 and 2009 have profoundly influenced my understanding of the dynamics of love and collectivist depersonalisation, and the whole approach of this essay.

'love' isn't love, but he must get himself somehow to believe it is. As Lillian Smith, who as a child faced the same forced choice, says, "It was the only way my world could be held together" (1963, 27). Lynchings were only the most lurid manifestation of the violent exclusion of black people that "not only divided the races but divided the white child's heart" (Smith 1963, 116). Like other well-to-do white southern children, Smith was largely brought up by a black nurse, but at twelve she knew – or, rather, she says, "I knew but I never believed it" – that her deep love for Aunt Chloe was "a childish thing which every normal child outgrows"; "I learned to cheapen with tears and sentimental talk of 'my old mammy' one of the profoundest relationships of my life. I learned the bitterest thing a child can learn: that the human relations I valued most were held cheap by the world I lived in" (1963, 18; cf. 112). She thereby also learned that there was "something Out There" – the collective racist world-order – "stronger" than her parents because they bowed to it, and "that people who talked of love and children did not mean it" (1963, 26–7).[11]

This is the violent meaning of the 'idolatry of collectivity'; individual human beings and the love they feel for each other are ultimately regarded as valueless, and indeed as threats to the collectivity that assigns individuals who are 'one of us' their value. If the dehumanisation of racism doesn't only oppress but also *attracts* us, this may be because its inherent depersonalisation of oppressors and oppressed alike appears to release us from the challenge of strictly personal contact and self-revelation where "I must make myself I to someone's Thou" (Gaita 2004, 338). Racism, then, would be a symptom of our anguished inability to stand our own humanity.

[11] See Wallace-Sanders 2008 on the violence done to *the black women* looking after white children who would then abandon them, while often forced to neglect or being forcibly separated from their own children.

References

Backström, J., 2007. *The Fear of Openness*. Åbo: Åbo Akademi University Press.
Backström, J., 2019. "Pre-truth life in post-truth times". *Nordic Wittgenstein Review*, Special Issue – Post-Truth: 97–130.
Backström, J., 2023. "Unbearable certainties". In: C. Eriksen et al. (eds.), *Philosophical Perspectives on Moral Certainty*. New York: Routledge, 76–97.
Backström, J. and Nykänen, H., 2016. "Collectivity, evil and the dynamics of moral value". *Journal of Evaluation in Clinical Practice* 22: 466–76.
Baldwin, J., 1965. *Going to Meet the Man*. London: Michael Joseph.
Baldwin, J., 2010. *The Cross of Redemption*. New York: Pantheon Books.
Bay, M., 2000. *The White Image in the Black Mind*. New York: Oxford University Press.
Bean, A., Hatch, J. and McNamara, B., 1996. *Inside the Minstrel Mask*. Hanover, NH: Wesleyan University Press.
Blassingame, J., 1979. *The Slave Community*. New York: Oxford University Press.
Butler, J., 2005. *Giving an Account of Oneself*. New York: Fordham University Press.
Cable, G. W., 1903. *The Negro Question*. New York: Charles Scribner's Sons.
Cox, O. C., 1959. *Caste, Class, and Race*. New York: Monthly Review Press.
Douglass, F., 2009. *Narrative of the Life of Frederick Douglass*. Cambridge, Mass.: Harvard University Press.
Du Bois, W. E. B., 1964. *Black Reconstruction in America*. Cleveland: Meridian.
Elkins, S., 1976. *Slavery*. Chicago: University of Chicago Press.
Ellison, R., 2003. *The Collected Essays of Ralph Ellison*. New York: Modern Library.
Fanon, F., 2008. *Black Skin, White Masks*. London: Pluto Press.
Fields, B. and Fields, K., 2012. *Racecraft*. London: Verso.
Fryer, P., 2018. *Staying Power*. London: Pluto Press.
Fox-Genovese, E. and Genovese, E., 2005. *The Mind of the Master Class*. Cambridge: Cambridge University Press.
Gaita, R., 1990. "Ethical Individuality". In: R. Gaita (ed.), *Value and Understanding*. London: Routledge, 118–148.
Gaita, R., 1994. "Critical Notice of D. Z. Phillips, *Interventions in Ethics*". *Philosophical Investigations* 17(4), 613–628.
Gaita, R., 2000. *A Common Humanity*. London: Routledge.
Gaita, R., 2004. *Good and Evil. An Absolute Conception*. Second Edition. London: Routledge.
Gaita, R., 2005. *The Philosopher's Dog*. London: Routledge.
Gaita, R., 2011. *After Romulus*. Melbourne: The Text Publishing Company.
Gaita, R., 2014. "Afterword – Anne Manne and Raimond Gaita in Conversation". In: C. Taylor and M. Graefe (eds.), *A Sense for Humanity*. Clayton: Monash University Publishing, 165–195.
Gaita, R., 2018. "Genocide". In: V. Harrison (ed.), *The History of Evil in the Early Twentieth Century: 1900–1950*. London: Routledge, 202–223.

Hartman, S., 2022. *Scenes of Subjection*. New York: W. W. Norton.
Hertzberg, L., 2022. "Absolutely Personal: A Countercurrent in Moral Philosophy". In: S. Aldrin Salskov et al (eds.), *Ethical Inquiries after Wittgenstein.* Cham: Springer, 105–121.
hooks, b., 2015. *Black Looks*. New York: Routledge.
Kolchin, P., 1987. *Unfree Labor*. Cambridge, Mass.: Harvard University Press.
Lhamon, W. T., 2003. *Jump Jim Crow*. Cambridge, Mass.: Harvard University Press.
Lomax, A., 1960. *The Folk Songs of North America*. New York: Doubleday & C:o.
Lott, E., 2013. *Love and Theft*. Oxford: Oxford University Press.
Memmi, A., 2000. *Racism*. Minneapolis: University of Minnesota Press.
Morrison, T., 1993. *Playing in the Dark*. New York: Vintage.
Nykänen, H., 2002. *The "I", the "You" and the Soul*. Åbo: Åbo Akademis förlag.
Nykänen, H., 2009. *Samvetet och det dolda: om kärlek och kollektivitet*. Stockholm: Dualis.
Nykänen, H., 2019. "This Thing with Philosophy". In: J. Backström et al (eds.), *Moral Foundations of Philosophy of Mind*. Basingstoke: Palgrave, 329–362.
Orwell, G., 1962. *The Road to Wigan Pier*. Harmondsworth: Penguin.
Roediger, D., 2007. *The Wages of Whiteness*. London: Verso.
Sartre, J.-P., 1995. *Anti-Semite and Jew*. New York: Schocken.
Scott, J. C., 1990. *Domination and the Arts of Resistance*. New Haven: Yale University Press.
Smith, L., 1963. *Killers of the Dream*. New York: Anchor Books.
Wallace-Sanders, K., 2008. *Mammy*. Ann Arbor: University of Michigan Press.
Weil, S., 2001. *Waiting for God*. New York: Harper Collins
Weil, S., 2005. "Human Personality". In: *Simone Weil: An Anthology*. London: Penguin, 69–98.

Biographical Note

Joel Backström, Ph.D., is researcher (Title of Docent) in philosophy at the University of Helsinki, where he also teaches philosophy. He is the author of *The Fear of Openness: An Essay on Friendship and the Roots of Morality* (ÅAUP, 2007) and co-editor of *Moral Foundations of Philosophy of Mind* (Palgrave, 2019) and of *Den okände von Wright* (SLS, 2023), a collection of G.H. von Wright's Swedish language writings in cultural criticism. Backström has published articles and book-chapters on ethics, on Wittgenstein and on the philosophical dimensions of Freud's thought, including contributions to *The Creation of Wittgenstein* (Bloomsbury 2023, ed. T. Wallgren), and to *Oxford Handbooks* on *Wittgenstein* (2011), *Philosophy and Psychoanalysis* (2018), and *Elizabeth Anscombe* (2022).