IRIS - Archivio Istituzionale dell'Università degli Studi di Sassari

Race, again: how face recognition technology reinforces racial discrimination

Questa è la versione Post print del seguente articolo:

Original
Race, again: how face recognition technology reinforces racial discrimination / Bacchini, F., Lorusso, L.. - In: JOURNAL OF INFORMATION, COMMUNICATION & ETHICS IN SOCIETY. - ISSN 1477-996X. - 17:3(2019), pp. 321-335. [10.1108/JICES-05-2018-0050]

Availability:
This version is available at: 11388/219924 since: 2022-06-09T13:25:26Z

Publisher:

Published
DOI:10.1108/JICES-05-2018-0050

Terms of use:

Chiunque può accedere liberamente al full text dei lavori resi disponibili come “Open Access”.

Publisher copyright

note finali coverpage

(Article begins on next page)

21 November 2026

Fabio Bacchini
Ludovica Lorusso

Race, again. How face recognition technology reinforces racial discrimination

Postprint. Please cite version in *Journal of Information, Communication and Ethics in Society*, 17, 3, 2019, pp. 321-335.

ABSTRACT

Purpose: This study aims to explore whether face recognition technology - as it is intensely employed by state and local police departments and law enforcement agencies - is racism-free or, on the contrary, is affected by racial biases and/or racist prejudices, thus reinforcing overall racial discrimination.

Design/Methodology/Approach: The study investigates the causal pathways through which face recognition technology may reinforce the racial disproportion in enforcement; it also inquires whether it further discriminates black people by making them experience more racial discrimination, and self-identify more decisively as black - two conditions that are shown to be harmful in various respects.

Findings: This study shows that face recognition technology, as it is produced, implemented and used in Western societies, reinforces existing racial disparities in stop, investigation, arrest and incarceration rates due to racist prejudices, and even contributes to strengthen the unhealthy effects of racism on historically disadvantaged racial groups, like black people.

Practical implications: The findings hope to make law enforcement agencies and software companies aware that they must take adequate action against the racially discriminative effects of the use of face recognition technology.

Social implications: This paper highlights that no implementation of an allegedly racism-free biometric technology is safe from the risk of racially discriminating, simply because each implementation leans against our society, which is affected by racism in many persisting ways.

Originality/Value: While the ethical survey of biometric technologies is traditionally framed in the discourse of universal rights, this study explores an issue that has not been deeply scrutinised so far, that is, how face recognition technology differently affects distinct racial groups, and how it contributes to racial discrimination.

Keywords: biometrics; face recognition; race; racial discrimination; racial identity

1. Introduction

The use of face recognition technology is becoming increasingly common all around the world. As algorithms get more reliable every year, more and more law enforcement agencies make use of biometric technologies that mimic human abilities to recognize faces. The benefits of face recognition technology are real, and law enforcement officers are importantly assisted by it in their effort to protect citizens and guarantee social security.

It would be naive, however, to think that the spread of this technology brings with it no unfortunate side effect. And in fact civil libertarians have not hesitated to denounce the risks for our privacy implicated by such a diffusion of automated face recognition systems. A hot debate exists today over how face recognition technology impacts privacy and civil liberties, and whether such important issues as civil rights, transparency and accountability are threatened by its dissemination, and to what degree (e.g. Bowyer, 2004; Cammozzo, 2011; Finn *et al.*, 2013; Senior and Pankanti, 2011).

These concerns are of the utmost importance, and it is not our intention to portray them as less significant than scholars agree about. Rather, our goal is trying to highlight a different sort of worries that so far have remained invisible as an effect of the general tendency to focus on aforementioned questions. When we deal with biometric technologies - and with face recognition technology in particular - we are only concerned with privacy issues. But framing the debate in the discourse of universal rights obscures the topography of the different degrees of heaviness of the hand of surveillance for different demographics. “Through that lens [the lens of the discourse of universal rights], government surveillance is seen as inflicting its harms on everyone” - Gellman and Adler-Bell (2017) say - but “surveillance is not at all the same thing at higher and lower elevations on the contour map of privilege.”

Our paper aims to show that face recognition technology leans against a society deeply affected by racial disparities produced by past and present racist attitudes, and in consequence confirms and even reinforces racial discrimination. We will focus on some causal pathways through which this happens; and we will conclude that - unless adequate countermeasures are adopted - high is the risk that face recognition technology further afflicts historically disadvantaged racial groups like black people - not just by exposing them to a higher probability to be incorrectly marked as suspects in crimes, but also by making them more stressed, more hopeless about the power of white racism and - above all - more unhealthy.

2. Racial disproportion in stop, investigation, arrest and incarceration rates as an effect of racism

In Western countries, black people are more likely subject to be stopped and identified by a police officer, investigated, arrested, incarcerated and even sentenced at all courts than white people. In 2009 in the USA the incarceration rate was 4,7 per cent for African Americans, and only 0,7 per cent for non-Hispanic whites (West 2010). In 2014, African Americans represented 5.4 per cent of Minnesota’s population but 24.5 per cent of those arrested. In the same year, the ratio of African American arrest rates to population share was 2:1 in Hawaii, Michigan and Virginia; 3:1 in Arizona, Pennsylvania, Los Angeles County and San Diego County; and 5:1 in Minnesota (Garvie *et al.*, 2016). This disproportion is not confined to the USA. In 2009-10 the UK overall population was 88,6 per cent white and only 2,7 per cent black, but the percentage of persons arrested and sentenced at all courts for indictable offences was 79,6 and 72 per cent, respectively, for whites, and 8 e 13,7 per cent, respectively, for blacks (UK Ministry of Justice, 2011).

One may just reply that these data are explained by the fact that black citizens commit more crimes. Yet the issue is far more complicated than this. As early as 1999, a report by the New York State Attorney General (Spitzer, 1999) highlighted that blacks and Hispanics represented 51 and 33 per cent of the people “stopped and frisked” by the New York City Police Department during the covered period of January 1, 1998 through March 31, 1999, despite being only 26 and 24 per cent of the city population. To reject the hypothesis that “higher crime rates in minority communities fully explain the higher rate at which minorities are stopped in New York City”, the report compared the stop rates to the number of crimes committed by members of each ethnic group, and concluded that “after accounting for the effect of differing crime rates, during the covered period, blacks were “stopped” 23 per cent more often than whites, across all crime categories. In addition, after accounting for the effect of differing crime rates, Hispanics were “stopped” 39 per cent more often than whites across crime categories” (*ibidem*, pp. ix-x). A subsequent 2013 report acknowledged that, although “the number of stops conducted by the New York City Police Department has grown dramatically over the last fourteen years, from 69,000 stops in 2000 to more than 685,000 stops in 2011”, the stop rates reflect even more significant racial disparities, since “stops of black and Hispanic individuals account for not only the majority of stops each year, but also the majority of the increase of stops

over the past fourteen years" (Schneiderman, 2013, p. 5). Indeed 88 per cent of the 4.4 million stops in New York City between January 2004 and June 2012 "resulted in no further action — meaning a vast majority of those stopped were doing nothing wrong". Yet in 83 per cent of cases, the person stopped was black or Hispanic, even though the two groups accounted for just over half the population (The New York Times, 2013). It is also interesting to note that, although the association among annual marijuana use and black race in the USA has been proved by many studies not higher - and even, in some case, significantly lower - than the same association with respect to white race (e.g. Johnston *et al.*, 2010), officers in New York City stop blacks on suspicion of marijuana possession at a rate of 14.83 per 1,000 population, while whites are stopped only 1.96 times per 1,000 population (Geller and Fagan, 2010, p. 608). In a similar way, the ratio among police stops and seizures of illicit goods in New York City in 2004-2012 was 27 for white and 143 for black people (Torres, 2015); and black people in the USA are up to 2.5 times more likely to be targeted by police surveillance than members of any other race (Garvie and Frankle, 2016). In short, overall enforcement - stops, arrests, and incarcerations - is higher for black people across crime types, also in cases in which this demographic group has lower levels of criminal involvement. It is easy to agree that we can only explain these data by referring to the racist prejudices of police officers and judges. True, in certain cases black people get more arrested or sentenced just because they commit a specific crime more often than members of other races. But we can explain also *this latter kind of fact* - the fact that black people commit a specific crime more frequently than white people - by appealing to the many negative as well as positive effects (economical, social, cultural, psychological) of past and present racism on the life conditions and personal choices of people belonging to different racial groups. The social force of white racism produced racial disparities in wealth throughout history, and demographics that are worse off are more likely to opt for criminal behaviour. Thus the wide range of different forms racism takes in our societies can account for *all these data*.

3. How face recognition reinforces the racial disproportion in enforcement

Our claim is that intense use of face recognition by state and local police departments contributes to strengthen these racial patterns via several distinct causal pathways:

1. Black people are overrepresented in many of the databases faces are routinely searched against. In fact several law enforcement agencies employ databases, or networks of databases, enrolling mug shots of individual arrested; and - as we have said - black people are disproportionately implicated in this category. Since a face recognition system can only "find" people who are in its database, African Americans are more often found as a match for a given suspect, producing a disproportionate number of both *true* and *false accept*. This in turn entails that black people are more often stopped, investigated, arrested, incarcerated and sentenced *as a consequence of face recognition technology*. In other words, the use of face recognition systems by law enforcement *strengthens and perpetuates* the racially disproportioned pattern we have described, contributing to make our societies even more unfair. Two clarifications are in order. First, a racial group can be negatively discriminated also by making it subject to a disproportionate number of *true accept*. In fact this means that members of other racial groups are more likely to get off scott-free when they break the law or commit a crime. We may discuss whether this is a *negative* discrimination, or just a *positive* discrimination with respect to white people; nonetheless there is no doubt that it is *racial discrimination*. Second, many agencies do employ mug shots databases with no limits or rules to limit which mug shots are enrolled based on the seriousness of the offence. Moreover, normally mug shots are not affirmatively "scrubbed" by police to eliminate no-charge arrests or not-guilty verdicts (Garvie *et al.*, 2016). This makes it even more probable for a black individual to be lifetime in a mug shot database as a mere effect of her being black - that is, for example, as an effect of her once having been unjustifiedly arrested by a zealous, racist white police officer while participating in a peaceful protest against the mistreatment of black people. Entering lifetime in a mug shots database - no matter you were never charged, had charges dropped or dismissed, or were found innocent - simply makes you more likely investigated as a suspect for other crimes, and eventually found (falsely) guilty of them. This is another positive feedback face recognition technology is causally responsible of, specifically afflicting law-abiding people belonging to socially disadvantaged racial groups.

2. Individuals with darker skin colours may be more difficult to identify because of the importance of colour contrast to characterise facial features, as some experts recently claimed (Garvie *et al.*, 2016, p. 88). Objective augmented difficulties in recognising persons belonging to darker skin colour racial groups would entail that these people be more likely *unjustifiedly* investigated, stopped or arrested, as well as *incorrectly* found as a match for a given suspect, independently of (1).

3. Many software display disturbing differences in accuracy across race. Klare *et al*. (2012) tested the recognition accuracies of six different face recognition algorithms (three commercial, two non-trainable, and one trainable), and found that both commercial and the non-trainable algorithms consistently have lower matching accuracies on black people (as well as on females and age group 18 to 30). The main reason seems to be racial disproportion in training: the same study discovered that the matching accuracy on black people can be improved by training exclusively on that racial group. Similarly, Buolamwini and Gebru (2018) examined three commercial gender classification systems and found that, while the maximum error rate for lighter-skinned males was 0.8 per cent, the error rates for darker-skinned females were up to 34.7 per cent; at the same time, they ascertained that two main facial analysis benchmarks (IJB-A and Adience) revealed overrepresentation of lighter males, underrepresentation of darker females, and underrepresentation of darker individuals in general.[1] Indeed most software programs are produced in countries where software engineers are predominately Caucasian males (Breland, 2017). We may suppose that the "other race effect" notoriously affecting human face recognition (Meissner and Brigham, 2001; Kelly *et al.*, 2007) gets inadvertently transmitted by white software engineers dominating the technology sector to the face recognition algorithms they produce via flawed, racially biased construction and training. Codes are geared to focus on white faces, mostly tested on white subjects, and trained on data sets mostly composed of white faces. That the racial bias depends on who builds the algorithms is confirmed by a number of studies showing that, while Western algorithm recognize Caucasian faces more accurately than East Asian faces, East Asian algorithm recognize East Asian faces more accurately than Caucasian faces (e.g. Phillips *et al.*, 2011). However, sometimes the racial bias is the (inadvertent) consequence of an intentional design choice. Garvie *et al.* (2016) report the case of the 2014 Handbook for Users of the Pennsylvania Justice Network that, on instructing them on how to generate a three-dimensional model of a face by using software from a company called "Animetrics", invites to select one category from the following list: "Generic Male, Generic Female, Asian Male, Asian Female, Caucasian Male, Caucasian Female or Middle Eastern Male". It is evident that African Americans (11.7% of Pennsylvanians) - along with other racial and ethnic groups - are completely ignored as an independent racial/ethnic category, thus creating the ideal conditions for a racial bias in accuracy. As an overall effect of these biases, black people are more likely to be *unjustifiedly* investigated, stopped or arrested, and *incorrectly* found as a match for a given suspect, independently of (1) and (2).

4. No accuracy test for racially biased error rates is regularly run by the companies or the law enforcement agencies using their software packages. No doubt that the racial biases in matching accuracy could be solved *ex post facto* if only such tests were run. Yet nobody seems interested in that this happens. No company or public agency ever publicly acknowledged the need to take corrective steps. Many jurisdictions also omit to disaggregate arrest rates along the lines of race and ethnicity.[2] We may even suppose that companies are implicitly dissuaded by the nature and public attitudes of their customers from the need to take action. Software developers may in good faith suppose that the main facial analysis benchmarks against which they test their algorithms are

[1] Buolamwini and Gebru (2018) introduce a new facial analysis dataset, the Pilot Parliaments Benchmark, which is balanced by gender and skin type. They chose to focus on skin type rather than on race and ethnicity because subjects' phenotypic features can vary widely within a racial or ethnic category. Also, racial and ethnic categories are not consistent across geographies and time. Since, in consequence, there are many skin types of individuals identifying as the same race, it seemed more appropriate to these scholars to use skin type rather than race to measure database diversity.

[2] More widely, in the USA "the facial-recognition algorithms used by police are not required to undergo public or independent testing to determine accuracy or check for bias before being deployed on everyday citizens" (Garvie and Frankle, 2016).

balanced by race and ethnicity - or, alternatively, by skin type - as well as by gender; but they are not (Buolamwini and Gebru, 2018). Notice that the *absence* of accuracy tests for racially biased error rates *is not an independent factor* that - along with (1), (2) and (3) - produces an increased probability for black people to be unjustifiedly investigated, stopped or arrested, or incorrectly found as a match for a given suspect. Indeed the *absence* of accuracy tests can produce that effect only if it is coupled with (3). Yet it is a further causal factor responsible for the effect - and it seems produced in turn by a typical lack of concern for racist attitudes and their effects in our societies that is morally blamable and - at least often - itself racist.

5. Human reviewers should be specifically trained to backstop inaccuracy, but normally they are not. Beyond racial biases, software programs are far from being highly reliable, and need to be double-checked by human users. Consider that many software pieces return the top few matches for a given suspect, no matter how bad the matches themselves are. In absence of a targeted training, there is a high risk to mistake the first (perhaps bad) match for an *indisputable* match needing no further scrutiny. Also consider that human reviewers (and even more so *untrained* human reviewers) are not only subjected to the "other race effect", but also - both consciously and unconsciously - to their own racial prejudices. Since most reviewers are Caucasian - and, in any case, not black - in Western societies, these latter factors turn out to penalise black people. For example, a white untrained reviewer may spontaneously apply greater scrutiny whenever the top positive match returned by the system is a white face, and omit to check that the results of face recognition searches are correct whenever the top positive match is a black face. Specialised training would be the only - though imperfect - solution, but very few law enforcement agencies seem aware of the problem. Garvie *et al.* (2016) even discovered that only eight over fifty-two agencies they found to use face recognition in the U.S. employed human gatekeepers to systematically review matches before forwarding them to officers. Human *trained* double-check is very important, however, as proved by a recent study: White *et al.* (2015) tested two groups of passport issuance staff, and a comparison group of untrained participants, on a task specifically designed to emulate workflow of passport issuance officers using face recognition software to detect fraudulent passport applications. While the first group of passport issuance officers had received limited training, the second group was specialist staff who had been purposely trained and included court-practicing forensic examiners with many years' experience in making face matching decisions. The results showed that the difference between untrained review staff and students was non-significant - both were incorrect *on over half of* all trials, thus producing misidentifications, by selecting the wrong identity, and misses by indicating that the target is not present in the array - while trained staff made substantially fewer errors, producing e.g. 20% fewer misidentifications. Even if an agency were persuaded to introduce regular training for its operators, however, it would encounter another difficulty: human reviewer training regimes are still at very early stages of development - probably because of the general lack of concern about inaccuracy.

As an overall effect of these five factors, black people are more often stopped, investigated, arrested, incarcerated and sentenced *as a consequence of face recognition technology*. The fact that a racial pattern pre-exists does not preempt face recognition technology from reinforcing it day after day, thus independently contributing to perpetuate it. Black people are more likely to be enrolled in face recognition systems, be subject to their processing and misidentified by them - with all the consequent troubles that white people more frequently escape. As U.S. House oversight committee ranking member Elijah Cummings effectively summed up in a congressional hearing on law enforcement's use of facial recognition software in March 2017, "If you're black, you're more likely to be subjected to this technology and the technology is more likely to be wrong" (Breland, 2017).

Day after day, people are becoming more and more aware of how face recognition technology negatively affects black people. Brian Brackeen, a black chief executive of a software company developing facial recognition services based in Miami, has categorically said that "there is really no "nice" way to acknowledge […] the potential dangers associated with current racial biases in face recognition", and is open in his "opposition to the use of the technology in law enforcement" (Brackeen, 2018). In UK Big Brother Watch – an independent non-profit organisation leading the protection of privacy and civil liberties – has recently stressed that "disproportionate misidentifications risk increasing the over-policing of ethnic minorities on the premise of technological 'objectivity'" (Big Brother Watch, 2018). And the media are increasingly bringing attention to real-world cases exemplifying how racial disproportion in the use of stop

and search powers by law enforcement agencies – as well as racial discrimination at all stages of criminal justice – is reinforced by the racial biases of face recognition technology (e.g. Berman and Lowery, 2015; Carlo, 2018; Dodd, 2018; Gayle, 2016; Murdock, 2017).

4. A further psychological price for black people

Consider that getting in trouble with the police and the law - even if you are innocent - negatively affects your life in a plethora of ways, from decreasing your professional opportunities and your average salary to causing psychological stress, which in turn notoriously produces numerous health impairments (e.g. Juster *et al.*, 2010). But this is not the end of the story. We must consider an additional side of the psychological price black people must pay.

First, black people are aware that they are more often stopped, investigated, arrested, incarcerated and sentenced in Western societies. They do remark - even perceptually - that, say, arrest and incarceration rates in the USA disproportionately implicate African Americans. They see that they get more frequently stopped than white people for questioning and inspection when coming through the border. And they interpret this disproportion as a consequence of the enduring presence of racism in our societies. They tend to think, at least on some occasions: "Again. A black person stopped by two white policemen on patrol. This is not happening by chance. This is happening because I am black". So not only black people get more often stopped, investigated, arrested, incarcerated and sentenced; they also *live these events* very differently with respect to white people, because they are inclined to interpret them as *racist episodes*. Since - as we have said - face recognition technology increases the probabilities that a black person - in virtue of her being black - gets in trouble with the police and the law, *it also increases the perceived racial discrimination that black people experience*. As we are going to explain, however, perceived racial discrimination is a further independent factor causing both social disadvantages and serious health impairments.

Second, the rapid spread of biometric technologies, and their intense use by police departments and law enforcement agencies - coupled with the enduring of the racial disproportion in the stop, arrest and incarceration rates - can be interpreted by black people as if biometric technologies *do see race* - contrary to the opposite claim that they (and specifically face recognition systems) "do not see race".[3] In any case, biometrics implements the idea that people *are their bodies*. Because people are identified thanks to their bodily traits, black people cannot think of themselves as persons whose identity is independent of their bodily traits. They cannot think of the physical traits that biometrics uses to identify them *as inessential* to define them as specific persons. But these physical traits are also those that mostly make them *black* rather than white or Asian: their DNA, first of all, but also their facial traits - lips, nose, eyes, hair, jawline - whose specific aspect appears so highly correlated to their skin colour. In consequence, black people are induced by the spread of biometric technologies, and particularly of face recognition, to heighten their self-identifying (a) as members of one race, rather than as members of any other human category (professional, religious, social, political, artistic, and so on); and (b) as members of the black race, rather than as members of any other racial group.

As we are going to point out, however, also decisively self-identifying as black people (in both senses (a) and (b)) in a society that is believed by the subject to be affected by racism towards black people, brings about social negative consequences as well as health impairments.

It is worth stressing that face recognition technology has a strong impact on self-identification. Many scholars in the last years have faced the topic of the relationship between biometric technology and people's identity and pointed out that biometric technology has a strong influence on identity, because it creates and uses body-based representations to identify persons (Alterman, 2003).

In particular, face recognition technology (1) reduces persons and their identity to their body, and (2) simplifies the plural nature of identity as well as its meanings and functions, being committed to the idea that there is *one and a single* identity that must remain the same across time and space. These two effects become very relevant when we consider that face recognition technology does not merely verify a pre-given identity, but also actively contributes to *creating and establishing an image of identity*, hence deeply transforming the subject who adopts that particular image (Bhabha, 1994; van der Ploeg, 1999, 2009). Because of face recognition technology's racial biases, these effects on people's identity may have profound consequences

[3] See the Seattle Police Department's Booking Photo Comparison System FAQs, Document p. 009377, where it is said that the system "does not see race, sex, orientation or age. The software is matching distance and patterns only, not skin color, age or sex of an individual" (Garvie *et al.*, 2016, p. 53).

on our societies, contributing to enhancing the racial boundaries between different racial groups as well as the *alienation* of the individuals belonging to some of these groups.
The first effect concerns the reduction of a person's identity to her body – that is, the reduction of the *who* to the *what*. This effect, which has been largely studied in social sciences and psychology, can have a different impact on different populations given that attributing a *biological foundation* of one's identity has not the same impact on different kinds of identities. Imagine a social context like the contemporary U.S., where blacks have been and still are discriminated because of their visible traits. For them, interpreting their racial identity as a social construct with no biological basis is of fundamental importance to perceive themselves free of racist stereotype threats and most importantly free to be *who* they want to be (Good *et al.*, 2007). All the worse racist prejudices are in fact founded on biological conceptions of race, as the literature in social psychology clearly shows (Williams and Eberhardt, 2008). Face recognition technology confirms black people that they *are* the observable constitution of their organism, that is, their phenotypic traits; that they are, in particular, the *look* of their face and its specific physical features; and, of course, that they are their skin colour. True, the same happens to white people; but the overall effect is very different. Indeed the reduction of one's identity to the body cannot be considered to have the same effects on people belonging to different populations. Urging white people to identify themselves with their body, facial traits, hair and skin colour means making them convinced that their essence coincides with those very visible features that cause them to be racially privileged in Western societies. This in turn produces an increase in their self-esteem, self-confidence and tenacity. In other words, white people end up to exploit even better their advantaged position in our societies - while members of racially disadvantaged groups will perform worse than as they would in the absence of face recognition technology. The racial gap in life conditions - as well as in resignation attitudes about them - rapidly expands as an effect of both processes.
The second effect concerns the oversimplification of the concept of identity. According to different studies and models developed in social psychology (e.g. Jones and McEwen, 2000) persons have multiple identities that are related to each other. In particular, following the 'social identity theory' and the 'self-categorisation theory' (Turner, 1984; Turner *et al.*, 1987), there exists a 'personal identity', which differentiates the unique self from all other selves, and a 'social identity', which is defined as the internalisation of often stereotypical, collective identifications. Group membership determines social identity and leads members to exaggerate the similarities within the in-group, and the differences between the in-group and out-group; for this reason group membership encourages members to discriminate against out-group members (Jenkins, 2008). Since social identity and self-categorisation strongly depend on the contexts in which people happen to live, and the contingencies with which they are faced every day, the context of surveillance should be in particular interpreted as potentially very dangerous because it can easily contribute to the reproduction and reinforcement of social divisions (Lyon, 2002, 2008). So face recognition technology boosts stereotypical categorial forms of social identity (*what kind* of person are you) and weakens the personal identity (*who* are you) of a person. Moreover, it makes the multiple dimensions of social identity (race, gender, sexual orientation, social class, religion, political beliefs, professions) collapse into just one, thus causing black people to merely self-identify as blacks and weakening - or even abolishing - all of the other social identities they have achieved during their life.
To sum up, face recognition technology heavily impacts self-identification, an in particular adversely affects individuals self-identifying as races negatively discriminated in our societies, thus reinforcing racial boundaries and racial discrimination itself. In the next section we shall show how both self-identification in discriminated categories and perceived discrimination are involved in complex causal pathways that may lead to various health impairments.

5. How the psychological price becomes physical

There is growing evidence that a major causal route exists from ongoing perception of racial discrimination to various health impairments via chronic stress and subsequent high allostatic load (Juster *et al.*, 2010; Seeman *et al.*, 2014). While homeostasis is the maintenance of the organism in a state in which all physiological parameters operate within normal values, allostasis refers to the process whereby the organism predictively change the parameters of its internal milieu to adapt to environmental stressors (Sterling, 2012). Allostatic load is defined as "the 'wear and tear' the body experiences when repeated allostatic responses are activated during stressful situations" (Juster *et al.*, 2010). Real or interpreted threats to homeostasis initiate the sympathetic-adrenal-medullary (SAM) axis release of catecholamines and the hypothalamic-pituitary-adrenal (HPA) axis secretion of glucocorticoids; "while adaptive acutely, chronic over-activation of SAM-

and HPA-axis products induce a 'domino effect' on interconnected biological systems that overcompensate and eventually collapse themselves, leaving the organism susceptible to stress-related diseases" (*ibidem*, p. 2). Many studies found higher values in biomarkers representing primary, secondary or tertiary outcomes in the allostatic load progression to reliably predict incident preterm birth and low birth weight, diabetes, cardiovascular disease, decline in physical functioning, decline in cognitive functioning, and mortality (Latendresse, 2009; Juster *et al.*, 2010; Seeman *et al.*, 2001). Indeed all these health impairments do shape along racial lines in Western societies and most typically in the USA, where black people have higher prevalence of most complex diseases (U.S. Department of Health and Human Services, 2010), even when they have a high socioeconomical status and education level - so we cannot merely explain the health differences by the racial differences in poverty, diet quality, access to health care and the similar. Nor should we be tempted to recur to a genetic explanation here: in fact, the hypothesis that differences in the risk of complex diseases among racial groups are largely due to genetic differences covarying with genetic ancestry appears highly problematic in the light of both current biological evidence and the theory of human genome evolution (Lorusso and Bacchini, 2015). The best explanation seems to be that systematic differences in ongoing levels of stress due to racial discrimination are causally responsible of the racial differences in the diseases rates (Geronimus et al., 2006). Black people are less healthy than white people not only because they are poorer - hence exposed to toxic substances, worse food and house conditions, lower levels of education - but also because they experience day after day more episodes of racism (which are sometimes almost undetectable, and yet unhealthy). In fact there is also a second biological mechanism connecting the psychological stress caused by racial discrimination, as a cause, and an unhealthy condition, as an effect: alteration in DNA methylation. Also known as 'epigenetic change', this is a process by which methyl groups are added to the DNA molecule and modify its activity. DNA methylation is activated by psychological stress and can be transmitted across generations, thus making the unhealthy effects of perceived racial discrimination temporarily *hereditary* (Kuzawa and Sweet, 2009; Thayer and Kuzawa, 2011; Lorusso, 2014).

We should not forget that experiencing ongoing racial prejudice is unhealthy also independently of the allostatic load and DNA methylation mechanisms. You might simply get depressed; or, you might lose access to health cares and good food as a consequence of your losing your job after reacting against a racist boss. Again, it is easy to imagine various causal pathways bringing from continuous perceived racial discrimination to different material and social disadvantages. For example, stress from discrimination contributes to interparental conflict, children witnessing interparental conflict are placed at heightened risk for emotional problems, and this impacts their early academic success and hence their professional luck (Lorusso and Bacchini, 2015).

On the contrary, it is not immediately evident why self-identifying as black may negatively affect your life. Yet self-identifying as black - coupled with the belief that black people are negatively discriminated in our societies - can produce feelings of resignation and powerlessness, which in turn adversely impact your life and the life of your offspring. For example, these feelings may lower teacher and mother expectations on black youth achievement outcomes, producing a disruptive effect on their actual achievements (Benner and Mistry, 2007). Such negative consequences are even more intense when self-identifying as black is lived rather as the acceptance of an inescapable condition than as a proudly combative choice - and this is exactly the kind of hint that self-identifying as black may take as an effect of the spread of face recognition and its capacity to "nail" people to their bodily features. Moreover, we know that racial self- and other-classification in Western societies are more fluid that commonly supposed. For example, Saperstein and Penner (2012) found that one in five Americans experienced at least one change in racial classification over a 19-year period. These shifts are not fortuitous, and rather reflect an individual's image of herself and her social success. Some of those who experience a decrease in their socioeconomic position see themselves as darker than before - and also as darker than others sharing their skin colour who had better luck. Similarly, some of those who experience an increase in their status are more likely to self-identify as white. These are not just changes in how people *visually* perceive themselves, of course. By modifying their racial classification, people seem to *explain* what happened to them. Self-identifying as black *as a result of one's misfortune* is tantamount to explaining it more or less this way: "I was doomed to fail. This is a racist society. I am black. Black people are discriminated. The result is only logical". We can agree that, at least sometimes, this is a *true* explanation of the facts. Once such an interpretation is elaborated and endorsed, however, the risk is high to feel deprived of one's self-confidence and energies. In other words, decisively self-identifying as black - paired with the belief that racism spreads its effects in our societies, and that people are nailed to their bodily features - can be seen as not only an effect, but also as a cause of a lower socioeconomic position. In a racist society, self-identifying as black and assigning high subjective importance to one's being black can

turn out to be - sad to say - self-fulfilling prophecies. One is likely to adopt more black-specific habits (diet, attire, hairstyle, slang, tastes, friendships) that may expose oneself to even more intense racial discrimination than before. Also, one is likely to become more able to detect even the subtlest racist episodes, and experience *more* racist discrimination than before - thus resulting more stressed, with all the pernicious effects on health that we have examined. Above all, people may develop hopelessness and resignation attitudes towards their disease risks, their low socioeconomic status and the racial discrimination they undergo, starting to omit to take adequate action. By "nailing" people to their bodily traits - as we have said - face recognition encourages black people not merely to self-identify as black, but to do so under a rigid, immovable, biologically founded concept of race. Williams and Eberhardt (2008, p. 1043) found that people viewing race as biologically derived, rather than as socially based, are more likely to "understand racial inequities as natural, unproblematic, and unlikely to change" as well as to be less motivated to seeking redress. It is easy to imagine how these attitudes can be further harmful and unhealthy in turn.

6. Conclusions

We have tried to show that - perhaps surprisingly - face recognition technology, as it is currently implemented, reinforces racial discrimination in Western societies. It is not just that face recognition *per se* increases the probabilities for black people - as opposed to white - to get stopped, investigated, arrested, incarcerated and sentenced, with all the resulting troubles; it also further negatively discriminates black people by making them experience more racial discrimination, and self-identify more decisively as black. This in turn negatively affects their lives in various ways.

A lesson we can learn is that no technological improvement is safe from the risk of contributing to racial discrimination if the social context in which it unfolds is heavily contaminated by racism and racial prejudice. Apparently there is no particular reason why face recognition technology should embody racism and help to keep alive disturbing racial patterns in the distribution of wealth and health. As the Seattle Police Departments effectively claimed, face recognition systems "do not see race". Or so it is easy to presume. Because software engineers, managers of software companies, police officers, chiefs of police, law enforcement agents, forensic scene investigators, government officials, attorneys, judges, law enforcement authorities, mayors, governors and, generally speaking, most citizens *do see race*, inevitably also face recognition technology - inasmuch as it is planned, produced, implemented and used by all these people - does see race, too, and is affected by racial prejudices. Moreover, even if face recognition technology were ideally racism-free in both the way it is designed and the way it is used, still it would get very soon polluted by the troublesome levels of racism affecting the many parts of the world it must come in contact with in order to operate. Consider, for example, the database problem: no matter how carefully a face recognition system is designed, and employed, to stay unencumbered by racist prejudices, it would fatally start strengthening racial discrimination - as we have shown - once it searches against databases enrolling mug shots of individual arrested, just because black people are disproportionately included in this category as an effect of many sorts of racist prejudice, past and present.

No implementation of face recognition technology, therefore, is safe from the risk of racially discriminating, simply because each implementation leans against our society which - sad to say - is affected by racism in many persisting ways. The first thing to do is to stop thinking that face recognition technology "does not see race", and start working hard to continuously monitor it and remove racist toxins from it. What we need is systematic and frequent mug shots databases scrubbing by the police; software training on data sets composed of faces representing all races; regular accuracy tests for racially biased error rates; specific training for human reviewers. Awareness is an obvious precondition. We cannot hope that any of these sub-goals is ever achieved unless we become definitively aware that face recognition technology is doomed to be racially biased - at least until racism is permanently erased. Of course, the most effective recipe for a racism-free face recognition technology is to struggle for a racism-free society. But it seems wise to couple the effort to remove the cause to that of containing and counterbalancing its harmful effects.

References

Alterman, A. (2003), "'A piece of yourself': ethical issues in biometric identification", *Ethics and Information Technology*, Vol. 5 No. 3, pp. 139-150.

Benner, A. D. and Mistry, R. S. (2007), "Congruence of mother and teacher educational expectations and low-income youth's academic competence", *Journal of Educational Psychology*, Vol. 99 No. 1, pp. 140-153.

Berman M. and Lowery, W. (2015), "The 12 key highlights from the DOJ's scathing Ferguson report", *The Washington Post*, 4 March, available at: https://www.washingtonpost.com/news/post-nation/wp/2015/03/04/the-12-key-highlights-from-the-dojs-scathing-ferguson-report/?noredirect=on&utm_term=.c54af3894c47 (accessed 13 September 2018).

Bhabha, H. (1994), *The Location of Culture*, Routledge, London.

Big Brother Watch (2018), *Face Off. The Lawless Growth of Facial Recognition in UK Policing*, May, available at: https://bigbrotherwatch.org.uk/wp-content/uploads/2018/05/Face-Off-final-digital-1.pdf (accessed 13 September 2018).

Bowyer, K. W. (2004), "Face recognition technology: security versus privacy", *IEEE Technology and Society Magazine*, Vol. 23 No. 1, pp. 9-19.

Brackeen, B. (2018), "Facial recognition software is not ready for use by law enforcement", *TechCrunch*, 25 June, available at: https://techcrunch.com/2018/06/25/facial-recognition-software-is-not-ready-for-use-by-law-enforcement/ (accessed 13 September 2018).

Breland, A. (2017), "How white engineers built racist code – and why it's dangerous for black people", *The Guardian*, 4 December, available at: https://www.theguardian.com/technology/2017/dec/04/racist-facial-recognition-white-coders-black-people-police (accessed 14 March 2018).

Buolamwini, J. and Gebru, T. (2018), "Gender shades: intersectional accuracy disparities in commercial gender classification", *Proceedings of Machine Learning Research*, Vol. 81 (Conference on Fairness, Accountability and Transparency), pp. 77-91.

Cammozzo, A. (2011), "Face recognition and privacy enhancing techniques", in Bissett, A., Ward Bynum, T., Light, A., Lauener, A. and Rogerson, S. (Eds.), *Proceedings of the Twelfth International Conference "The Social Impact of Social Computing - Ethicomp 2011"*, Sheffield Hallam University, Sheffield, pp. 101-109.

Carlo, S. (2018), "We've got to stop the Met Police's dangerously authoritarian facial recognition surveillance", *Metro*, 6 July, available at: https://metro.co.uk/2018/07/06/weve-got-to-stop-the-met-polices-dangerously-authoritarian-facial-recognition-surveillance-7687833/ (accessed 13 September 2018).

Dodd, V. (2018) "Facial recognition out, knife arches in at Notting Hill carnival", *The Guardian*, 24 August, available at: https://www.theguardian.com/culture/2018/aug/24/knife-arches-to-be-set-up-at-notting-hill-carnival (accessed 13 September 2018).

Finn, R. L., Wright, D. and Friedewald, M. (2013), "Seven types of privacy", in Gutwirth, S., Leenes, R., De Hert, P. and Poullet, Y. (Eds.), *European Data Protection: Coming of Age*, Springer, Dordrecht, pp. 3-32.

Garvie, C., Bedoya, A. M. and Frankle, J. (2016), *The Perpetual Line-up. Unregulated Police Face Recognition in America*, Center on Privacy and Technology, Georgetown University, October 18, available at: https://www.perpetuallineup.org (accessed 14 March 2018).

Garvie, C. and Frankle, J. (2016), "Facial-recognition software might have a racial bias problem", *The Atlantic*, 7 April, available at: https://www.theatlantic.com/technology/archive/2016/04/the-underlying-bias-of-facial-recognition-systems/476991/ (accessed 14 March 2018).

Gayle D. (2016), "Police officers call for Notting Hill carnival review after record arrests", *The Guardian*, 30 August, available at: https://www.theguardian.com/culture/2016/aug/30/notting-hill-carnival-arrests-hit-record-high (accessed 13 September 2018).

Geller, A. and Fagan, J. (2010), "Pot as pretext: marijuana, race, and the new disorder in New York City street policing", *Journal of Empirical Legal Studies*, Vol. 7 No. 4, pp. 591-633.

Gellman, B. and Adler-Bell, S. (2017), *The Disparate Impact of Surveillance*, The Century Foundation, 21 December, available at: https://tcf.org/content/report/disparate-impact-surveillance/ (accessed 11 March 2018).

Geronimus, A. T., Hicken, M., Keene, D. and Bound, J. (2006), ""Weathering" and age patterns of allostatic load scores among blacks and whites in the United States", *American Journal of Public Health*, Vol. 96 No. 5, pp. 826-833.

Good, C., Dweck, C. S. and Aronson, J. (2007), "Social identity, stereotype threat, and self-theories", in Fuligni, A. J. (Ed.), *Contesting Stereotypes and Creating Identities: Social Categories, Social Identities, and Educational Participation*, Russell Sage Foundation, New York, pp. 115-135.

Jenkins, R. (2008), *Social Identity*, Routledge, London.

Johnston, L. D., O'Malley, P. M., Bachman, J. G. and Schulenberg, J. E. (2010), *Monitoring the Future. National Survey Results on Drug Use, 1975-2009. Volume I: Secondary school students*, NIH Publication No. 10-7584, National Institute on Drug Abuse, Bethesda, MD.

Jones, S. R. and McEwen, M. K. (2000), "A conceptual model of multiple dimensions of identity", *Journal of College Student Development*, Vol. 41 No. 4, pp. 405-414.

Juster, R. P., McEwen, B. S. and Lupien, S. J. (2010), "Allostatic load biomarkers of chronic stress and impact on health and cognition", *Neuroscience & Biobehavioral Reviews*, Vol. 35 No. 1, 2e16.

Kelly, D. J., Quinn, P. C., Slater, A. M., Lee, K., Ge, L. and Pascalis, O. (2007), "The other-race effect develops during infancy: evidence of perceptual narrowing", *Psychological Science*, Vol. 18 No. 12, pp. 1084-1089.

Klare, B. F., Burge, M. J., Klontz, J. C., Bruegge, R. W. V. and Jain, A. K. (2012), "Face recognition performance: role of demographic information", IEEE Transactions on Information Forensics and Security, Vol. 7 No. 6, pp. 1789-1801.

Kuzawa, C. W. and Sweet, E. (2009), "Epigenetics and the embodiment of race: developmental origins of US racial disparities in cardiovascular health", *American Journal of Human Biology*, Vol. 21 No.1, pp. 2-15.

Latendresse, G. (2009), "The interaction between chronic stress and pregnancy: preterm birth from a biobehavioral perspective", *Journal of Midwifery & Women's Health*, Vol. 54 No. 1, pp. 8-17.

Lorusso, L., (2014), "The epigenetic hypothesis and the new biological role of self-identified racial categories", *Critical Philosophy of Race*, Vol. 2 No. 2, pp. 183-203.

Lorusso, L. and Bacchini, F. (2015), "A reconsideration of the role of self-identified races in epidemiology and biomedical research", *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, Vol. 52, pp. 56-64.

Lyon, D. (2002), "Everyday surveillance: personal data and social classifications", *Information, Communication & Society*, Vol. 5 No. 2, pp. 242-257.

Lyon, D. (2008), "Biometrics, identification and surveillance", *Bioethics*, Vol. 22 No. 9, pp. 499-508.

Meissner, C. A. and Brigham, J. C. (2001), "Thirty years of investigating the own-race bias in memory for faces: a meta-analytic review", *Psychology, Public Policy, and Law*, Vol. 7 No.1, pp. 3-35.

Murdock J. (2017), "Police to use 'racist' face-scanning tech at Notting Hill Carnival 2017", *International Business Times*, 16 August, available at: https://www.ibtimes.co.uk/police-use-racist-face-scanning-tech-notting-hill-carnival-2017-1635286 (accessed 13 September 2018).

Phillips, P. J., Jiang, F., Narvekar, A., Ayyad, J. and O'Toole, A. J. (2011), "An other-race effect for face recognition algorithms", ACM Transactions on Applied Perception (TAP), Vol. 8 No. 2, pp. 1-14.

Saperstein, A. and Penner, A. M. (2012), "Racial fluidity and inequality in the United States", *American Journal of Sociology*, Vo. 118 No. 3, pp. 676-727.

Schneiderman, E. T. (2013), *A Report on Arrests Arising from The New York City Police Department's Stop-And-Frisk Practices*, Office of the New York State Attorney General, Civil Rights Bureau, November, available at: https://ag.ny.gov/pdfs/OAG_REPORT_ON_SQF_PRACTICES_NOV_2013.pdf (accessed 18 April 2018).

Seeman, T. E., McEwen, B. S., Rowe, J. W. and Singer, B. H. (2001), "Allostatic load as a marker of cumulative biological risk: MacArthur studies of successful aging", *Proceedings of the National Academy of Sciences*, Vol. 98 No. 8, pp. 4770-4775.

Seeman, M., Merkin, S. S., Karlamangla, A., Koretz, B. and Seeman, T. (2014), "Social status and biological dysregulation: the "status syndrome" and allostatic load", *Social Science & Medicine*, Vol. 118, pp. 143-151.

Senior, A. W. and Pankanti, S. (2011), "Privacy protection and face recognition", in Li, Stan Z. and Jain, Anil K. (Eds.), *Handbook of Face Recognition*, Springer, London (pp. 671-691).

Spitzer, E. (1999), *The New York City Police Department's "Stop and Frisk" Practices.* Office of the New York State Attorney General, available at: https://ag.ny.gov/sites/default/files/pdfs/bureaus/civil_rights/stp_frsk.pdf (accessed 18 April 2018).

Sterling, P. (2012), "Allostasis: a model of predictive regulation", *Physiology & Behavior*, Vol. 106 No. 1, pp. 5-15.

Thayer, Z. M. and Kuzawa, C. W. (2011), "Biological memories of past environments", *Epigenetics*, Vol. 6 No. 7, pp. 798-803.

The New York Times (2013), "Racial discrimination in stop-and-frisk", 12 August, available at: https://www.nytimes.com/2013/08/13/opinion/racial-discrimination-in-stop-and-frisk.html (accessed 18 April 2018).

Torres, J. (2015), "Race/ethnicity and stop-and-frisk: past, present, future", *Sociology Compass*, Vol. 9 No. 11, pp. 931-939.

Turner, J. C. (1984), "Social categorization and the self-concept: a social cognitive theory of group behaviour", in Lawler, E. J. (Ed.), *Advances in Group Processes: Theory and Research*, Vol. 2, JAI Press, Greenwich, CT, pp. 77–122.

Turner, J. C., Hogg, M. A., Oakes, P. J., Reicher, S. D. and Wetherall, M. S. (1987), "Rediscovering the social group: a self categorization theory", Blackwell, Oxford.

UK Ministry of Justice (2011), *Statistics on Race and the Criminal Justice System 2010. A Ministry of Justice publication under Section 95 of the Criminal Justice Act 1991*, October, available at: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/219967/stats-race-cjs-2010.pdf (accessed 11 March 2018).

U.S. Department of Health and Human Services (2010), *Healthy People 2020*, available at: healthypeople.gov (accessed 14 March 2018).

van der Ploeg, I. (1999), "The illegal body: "Eurodac" and the politics of biometric identification", *Ethics and Information Technology*, Vol. 1 No. 4, pp. 295–302.

van der Ploeg, I. (2009), "Machine-readable bodies: biometrics, informatization and surveillance", in Mordini, E. and Green, M. (Eds.), *Identity, Security and Democracy*, IOS Press, Amsterdam, pp. 85-94.

West, H. C. (2010), "Prison inmates at midyear 2009 - statistical tables", Bureau of Justice Statistics, U.S. Department of Justice, NCJ 230113, June, available at: https://www.bjs.gov/content/pub/pdf/pim09st.pdf (accessed 14 March 2018).

White, D., Dunn, J. D., Schmid, A. C. and Kemp, R. I. (2015), "Error rates in users of automatic face recognition software", *PLoS One*, Vol. 10 No. 10, e0139827.

Williams, M. J. and Eberhardt, J. L. (2008), "Biological conceptions of race and the motivation to cross racial boundaries", *Journal of Personality and Social Psychology*, Vol. 94 No. 6, pp. 1033-1047.