Fairness, implicit bias testing and sports refereeing
An argument for why professional sports organisations ought to promote fairness in sport through testing referees for implicit biases
Petersen, Thomas Søbirk; Wichmann, Søren Sofus

Published in:
Journal of the Philosophy of Sport

DOI:
10.1080/00948705.2020.1866994

Publication date:
2021

Document Version
Peer reviewed version

Citation for published version (APA):

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain.
- You may freely distribute the URL identifying the publication in the public portal.

Take down policy
If you believe that this document breaches copyright please contact rucforsk@kb.dk providing details, and we will remove access to the work immediately and investigate your claim.
Fairness, implicit bias testing and sports refereeing

An argument for why professional sports organisations ought to promote fairness in sport through testing referees for implicit biases

ABSTRACT

Sports referees are not always as unbiased or impartial as they ideally should be. Studies have shown, for example, that in their decisions, referees seem to be biased against people of different race, gender or height or biased in favour of the home team. There is no doubt that such partiality work against official ideals of fairness and non-discrimination in sport. The problem with this is that being affected by implicit biases potentially causes the referees to make unfair decisions, with the result that some people are penalised or harmed disproportionately within the realm of sport. In this paper, we argue that sports organisations ought to require referees to undergo implicit bias testing as part of their mandatory training, in order to gain knowledge about biases and to take the proper counter-measures to combat such biases. Finally, we present and critically discuss four objections to our argument and conclude that none of them are plausible.

Keywords: fairness; refereeing; implicit bias testing; sports ethics
Fairness, implicit bias testing and sports refereeing

An argument for why professional sports organisations ought to promote fairness in sport, through testing referees for implicit biases

1. Introduction

There is scientific evidence for the claim that professional sports referees are not always as unbiased as they ideally should be. Studies have shown, for example, that in their decisions, referees seem to be biased against people of different race (see, e.g., Price and Wolfers [2010]; Tolany [2017]), gender (see, e.g., Souchon, Livingstone, and Maio [2013]) or height (Gift and Rodenberg [2014]) or biased in favour of the home team (see, e.g., Boyko, Boyko, and Boyko [2007]). When it comes to race, for example, studies have shown that African-American football players are twice as likely to be shown a red card than their non-African-American counterparts (Tolany 2017). Such decisions clearly go against official ideals of fairness and non-discrimination in sport – ideals that are explicitly accepted by several big sports organisations such as the international football association (FIFA 2019) or in the charter of the International Olympic Committee (IOC) which, as part of its fundamental principle of Olympism, states: ‘Every individual must have the possibility of practicing sport, without discrimination of any kind …’ (IOC 2019, 4).
However, there is no doubt that the very nature of their work makes sport referees highly susceptible to making biased judgments. Referees rely highly on intuitive reasoning in order to make the kind of decisions necessary in their work; very quick decisions that are often made under very high pressure from participants and crowds. Studies show how such judgments are vulnerable to implicit biases (Irwin and Real 2010) and unsurprisingly, as we have already hinted, studies show how sports referees are affected by implicit biases, primarily racial bias. The problem is, of course, that being affected by implicit biases potentially makes the referees make unfair decisions, with the result that some people are penalised disproportionately within the realm of sport. Besides these observations, such unfair decisions create frustration and harm among, for example, the athletes, coaches and spectators.

Apart from a few academic papers about the phenomenon of racist outcomes of sports referees’ decision-making, little has been done to highlight the problem and what to do about it. While sports organisations like FIFA (FIFA 2019) or government organisations like the Council of Europe have tried to combat biases such as racism in sport, the focus for both organisations, when it comes to the role of the referee, is expressed in the following: ‘… encourage referees to react appropriately where fans or athletes engage in racist behaviour by imposing adequate measures and sanctions, including by suspending competitions’ (Council of Europe 2009). Not a word about the possible racist behaviour of the referees themselves – a kind of behaviour that can be just as discriminatory and morally problematic as that of the fans or athletes. However, one possible and unexplored path to reduce biased refereeing is to have referees undergo implicit bias testing, in order to gain knowledge about biases and to take the proper countermeasures to combat such biases.
There is substantial evidence that tests such as Implicit Association Tests can be an effective tool in selection and education efforts to reduce the impact of implicit bias (Irwin and Real 2010; Kurdi and Banaji 2017). Given that such tests are in fact effective and practically feasible, the truth of which will be clarified in Section 2, this paper argues that sports referees should be tested for relevant implicit biases as part of their education and/or selection process. Before presenting the structure of the article, a comment about the scope seems appropriate. Throughout the paper we focus on the idea that referees should be tested for implicit bias. However, it may be just as relevant to test those who recruit, train, manage and evaluate the profession as what goes on in the selection process could also be a potential source of, for example, racial bias that indirectly causes racial biases among referees. However, we stick to our focus on testing the referees for two reasons. First, because there, contrary to the biases of sports referees (and for example, a jury member of a court trial) (see, e.g., Bennett [2010] for data supporting the observation that the selection of jury members are racially biased), is little scientific evidence that the process of selecting sports referees is influenced by bias, but also for stylistic reasons.

The structure of the paper is the following. Section 2 briefly describes the state and nature of testing for implicit cognitive biases and how they can be effective in changing biases. In Section 3, we present the argument for why referees should be tested for implicit bias. Section 4 critically discusses possible objections to the argument. Section 5 sums up the findings.

Before we move on, a few primary clarifications. First, we have chosen to focus on ‘implicit biases’ instead of explicit biases as explicit, for example, racial bias is a lot easier
to detect and therefore also easier to address. Second, one question that we need to answer is what kind of proposal we are making. Is it a moral obligation for the individual referee, or is it a systemic rule-based proposal that sports federations should test referees and that the tests should be mandatory? We think the latter, because it would be more efficient to combat racial bias, for example, if implicit bias testing were not reliant on the moral conscience of the individual referee. An individual who might not have explicit racist attitudes, and therefore does not consider that their behaviour could have a racist outcome, might behave in a way that produces discriminatory outcomes such as those mentioned in the first paragraph of this article. We also work with the latter strategy as, from a moral point of view, it is more controversial and would probably give rise to stronger objections (of which four will be clarified and critically discussed in Section 4), compared to just allowing and supporting the idea that referees take an implicit bias test. Thus, the following will discuss such a scenario, in which, in order to be eligible as a referee, potential and existing referees would be required to undergo implicit bias testing as part of their mandatory training.

2. Testing for implicit bias

Before we present our argument, it is appropriate to mark out the field concerning implicit bias testing. The following is a short section on the nature of implicit cognitive biases and how they can be revealed. Implicit cognitive biases are automatic associations which influence action without people’s reflective awareness of them. Because they are often unavailable to the introspective awareness of those who harbour them, it is possible for a
person to be explicitly unbiased, and yet be influenced by implicit biases, as these are activated involuntarily without awareness or intentional control by the individual (Brownstein and Saul 2016; Holroyd 2015; Staats and Patton 2013). Thus, a sports referee can have all the best intentions of being unbiased, and yet make calls that are implicitly biased. However, by taking an Implicit Association Test (IAT) or the like, these implicit biases can be revealed (Brownstein and Saul 2016; Staats and Patton 2013). Although testing positive for an implicit bias does not guarantee that a person will also act in a biased way, it is possible to make predictions of the likelihood of future behaviour from bias testing that is significantly better than random guesswork (Jost et al. 2009; Kang et al. 2011; Kurdi and Banaji 2017; Staats and Patton 2013). By this researchers like the aforementioned mean that if a person is found to be significantly implicitly biased towards, for example, a group of people, there is also an increased likelihood that this bias will be reflected in their behaviour.

More than two decades of research in cognitive psychology has shown how humans are to a very large extent implicitly biased (Banaji and Greenwald 2016; Holroyd, and Kelly 2016; Staats and Patton 2013). Implicit biases have been shown to affect judgments and behaviour in parts of society where ideals such as fairness and unbiasedness are thought to be integral – most prominently, in the criminal justice system, where a wide range of professionals have been shown to be implicitly biased, for example so-called ‘shooter bias’ in police officers (Correll et al. 2002), confirmation bias in fingerprint experts (Dror, Charlton, and Pron 2006; Kassin, Dror, and Kukucka 2013), racial bias in jury members (Kang and Lane 2010; Levinson, Cai, and Young 2010; Roberts 2011) and trial judges (Bennett 2016; Rachlinski et al. 2008).
Implicit biases are automatically acquired as by-products of our common culture, and although they are difficult to suppress and control, it is not impossible (Madva 2016; Rees 2016). However, an important condition for making that change is to make individuals aware of their own implicit biases. The most common method for this is the previously mentioned Implicit Association Test (see: https://implicit.harvard.edu/implicit/). With the present tests available, it is controversial whether tests such as the IAT can be used as a tool to select individuals, while using the test as an educational tool to become aware of one’s own hidden biases is much less controversial (Greely 2013; Kang et al. 2011; Mitchell and Tetlock 2017; Project Implicit 2019; Roberts 2011). Also, there are trajectories in the research which show that the tests could be improved to procure reliable information about individuals (Greenwald and Sriram 2010) Kurdi and Banaji 2017). The suggestion put forth in this paper is that sports organisations should use implicit bias testing as a tool to make their employed referees less biased, by using the information about the individual judges to improve anti-bias initiatives and education. Research on jury members and healthcare professionals, has shown how taking an Implicit Association Test can be a useful tool in learning to counter one’s own personal implicit biases (Roberts 2011). Furthermore, the advantages of this approach, over using implicit bias testing as a selection tool, where for example potential referees were to be excluded if they are above some level of implicit bias against a certain group are at least three-fold: First, the reliability of such individual use is, as mentioned above, very controversial. Second, it would also be controversial to deselect referees purely on the basis of an implicit bias test if their behaviour as referees does not reflect those biases. Third, there may be a lack of skilled referees if they all turn
out to be too biased. Research on jury members and healthcare professionals has shown how taking an Implicit Association Test can be a useful tool in learning to counter one’s own personal implicit biases (Roberts 2011).

3. The argument

Our argument proceeds in four steps, comprising three premises and a conclusion.

**Premise 1**: If sports organisations can promote fairness, provided no significant sacrifices are required, they ought to do it

**Premise 2**: Sports organisations can promote fairness in sport if they make it mandatory that potential and existing referees undergo implicit bias testing.

**Premise 3**: For sports organisations to have mandatory implicit bias testing of their referees would only require small sacrifices.

**Conclusion**: Sports organisations ought to make referees undergo mandatory implicit bias testing.

Some aspects of the argument require clarification. Concerning premise 1, we take for granted that sports organisations, and especially professional sports organisations like FIFA, who pay their referees and where huge sums of money are involved in the sport, have the ability to promote fairness in sport through paying for implicit bias testing of their referees and make sure that they actually use the tests. In addition, if testing for implicit
bias can improve the quality of refereeing, without significant sacrifices in terms of money or well-being, then the testing in question will, all else being equal, be very beneficial for the whole sports community and society in general. By providing more fairness through implicit bias testing, professional sports could be a better driver for the fair treatment of, for example, minorities. Although it is necessary to know in detail how the policy of mandatory implicit bias testing should be implemented and managed in order to evaluate our suggestions, we trust that all this can be done in a morally defensible manner. For example, by ensuring that the results of the tests, in order to protect anonymity and privacy are only revealed to the referee being tested and his or her supervisor and that the results cannot be traced by any third party. Finally, it is worth being aware that we are not claiming that implicit bias testing is a technological fix for bringing about perfect fairness in sport. Our suggestion is only that implicit bias testing is possible tool that can be used to improve fairness in sport.4 Concerning the plausibility of premise 2, which is an empirical premise, we rely on the material presented in Section 2.

Let us now turn to some possible objections to the argument that we have presented in this section.

4. Objections

Although they differ in many ways, all of the four following objections to our argument strike at premise 3, by claiming that there are costs, connected with testing sports referees for implicit bias, that are too great a sacrifice. One obvious cost that we will not spend much of our energy on, is the financial cost of testing referees. We will not do so for two
reasons. One is that if the tests used are anything like the present IAT, then they are available free online, in which case this will be a minor concern. Secondly, if tests are extended to using, for example, neuroimaging, they may indeed be quite costly;\textsuperscript{5} however, considering that professional sport is a billion dollar industry, we believe it is plausible that the money could be found to improve the industry in a way that fits with the goal of many sports organisations, namely that the sports results of this industry, are as fair as possible.

\textit{Objection 1: Trust and refereeing}

One reason why testing referees for implicit cognitive bias could be wrong could be that it incurs too great a sacrifice on the individual referee, because it violates their professional agency. The objection would be based on a claim that part of being a skilled professional referee is to be able to make decisions that are unbiased and that this agency is respected by employers and other parties involved. Even if one rules out that professional agency has intrinsic value such diminishing of agency could, nevertheless, be problematic for at least two reasons. One is that it is unpleasant for the referees, and the other is that public distrust can make the referees distrust their judgment, making them less good at their job.

First, we will deal with the unpleasantness aspect. It is well known that discovering that one is implicitly biased, especially in the case of biases that are socially unacceptable, for example racism or sexism, can be unpleasant if those implicit biases are not aligned with one’s explicit beliefs. The creators of the IAT even warn the reader against taking the
test because the results may reveal unpleasant surprises; there is a risk that the result ‘causes distress, even sadness, because it undermines the image we have of ourselves as largely fair-minded and egalitarian’ (Banaji and Greenwald 2016, 69). The question is whether such ‘awareness discomfort’ (Munch-Jurisic 2020) is too much of a personal sacrifice for the individual referee. However, there are good reasons to answer this question in the negative. Avoiding discomfort for the biased person has already been criticized for putting too much emphasis on the well-being of the biased person, at the expense of the person who is being exposed to biased treatment, and because it may miss out on the potential for change that comes with discomfort (Applebaum 2017; Bailey 2017). Being treated according to a bias is often worse than having a bias exposed, and if there is a disproportionality of harm between the awareness discomfort of the biased person and the discomfort experienced by the victim of biased treatment, then that disproportionality will be increased by the fact that the biased person is a referee, who has the power to influence the result of a game and in some situations to stall or destroy the career of the athlete.

Second, what if the awareness of a bias makes the referee distrust themselves to a degree where it makes their ruling more incorrect than before the testing? For example, the referee may begin to overcompensate by becoming biased towards other athletes that the referee was not biased against to begin with. If this is the case, and everything else is equal, such behaviour could be considered devastating for our argument. But we do not think that such possible scenarios would necessarily function as a knock-down argument to our argument. First, it all depends on what the overcompensation amounts to. Imagine that a referee, after having been tested for implicit biases has stopped being severely biased
towards Afro-American athletes, but instead has begun to be slightly biases against Caucasians. This scenario would not be a problem if we want to promote fairness, as less biased behaviour is better than more biased behaviour. On the other hand, if an implicit bias test results in more unfairness this would make our argument flawed. Second, however, and as we refer to in Section 2, such behaviour is not what we can expect according to the research in implicit bias testing, quite the contrary. It has been shown, however, that public officials such as trial judges perform better, if they have a high degree of public and personal trust in themselves (Peršak 2014). But such observations could easily speak in favour of our argument. For there are good reasons to believe that the public trust in referees could be increased if sports organisations did use implicit bias testing, thereby signalling that they do what they can to combat biases among referees. It should be mentioned that there are also scholars who argue that creating awareness of implicit bias is an important tool in making society more just (Applebaum 2017; Bailey 2017; Devine 2005). Being confident and comfortable makes for a good working environment for referees and most other professions. On the other hand, taking a more dialectical approach to one's own work also has its merits. Moreover, implicit bias testing of referees may also create more trust in refereeing and referees for spectators and sponsors of big sports events.

To conclude this section, there are no good moral reasons why the awareness discomfort of implicitly biased referees should count more than the risk of having athletes be treated unfair by those referees. Our response to the objection that the discomfort of having one’s implicit biases exposed is unacceptable because it may reduce the quality of the refereeing is an empirical question which calls for more research.
Objection 2: The rules of the game

Another line of objections to testing referees for implicit cognitive bias shifts the focus from the inner lives of the referees and others who perceive them, to the rules and principles that guide sports. By testing referees for implicit cognitive bias, something important to the nature of sports is sacrificed in order to produce more fairness. This approach can be divided into two separate yet related objections. One is that testing referees for implicit cognitive bias is wrong because trusting the rulings of the referee is an integral and important part of sports which will be lost if they are tested. The second is that, by testing for implicit bias, part of the judge’s responsibility is dispersed into technology and the experts that produce it, thus removing some of the authority of the referee. Both these types of objections have been raised against the use of implicit bias testing of judges in criminal trials (Kahn 2017).

In the first case, in many sports the fallibility of the referee is an integral part of the game. Soccer referees make mistakes, even in deciding Video Assistant Referee (VAR) cases (Collins 2012 and 2019), and part of the joy-filled anguish of being a soccer fan is discussing the quality of refereeing. What is important here is that in cases of, for example, implicit race bias, which is a problem in soccer, it is not individual random mistakes made by the referee. Clearly, it is not supposed to be part of the joy of watching soccer games to see African-American players be systematically yellow or read-carded more than players of other races, at least not according to FIFA. Mistakes and imperfections are part of the
game, but there is a crucial difference between individual, random mistakes by referees, and implicit biases directing those mistakes at ethnic or other groups.

The second objection, that implicit bias testing is an undue outsourcing of referee responsibility to those who produce implicit bias tests, is clearly a matter of policy. How much help should referees have? We will argue that in sports the use of, for example, VAR and Hawk-Eye in tennis (see, e.g., Collins and Evans [2018] for technological details of how Hawk-Eye works), also implies a dispersion of responsibility to the VAR referees and the camera operators. But if we accept the use of a technology like VAR and Hawk-Eye within certain limits and understanding, then why not accept that referees can improve their performance through implicit bias testing? We are well aware of the fact, that despite their proven positive impact on refereeing effectiveness, criticism of technologies such as VAR has mounted in some sports. In soccer, for example, spectators and journalists have criticized the VAR system on the grounds that it negatively affects both the flow of the game and the referee’s authority (i.e., the human authority on the field) and undermines important human elements such as verbal communication and fallibility (Holt 2019; Wilson 2018, ). But despite these voices there seems to be consensus on the verdict that the use of VAR and Hawk-Eye is here to stay, and that it is still the referee who is left on the court with the decision.

**Objection 3: Biases may be correct**

A third way in which testing referees can be said to carry disproportionate sacrifices is by claiming that eliminating them has a high epistemic cost. This objection is based on the observation that biases and heuristics are an integral part of human reasoning (Brownstein
and Saul 2016; Jost et al. 2009; Kahneman 2011) and stereotypical implicit biases might have a kernel of truth to them (Allport 1979) because they reflect the society in which they exist. Thus, some biases can serve as some sort of a statistical baseline for making predictions about the behaviour of others (Gendler 2011; Tetlock et al. 2000). Or, alternatively, framed in the present context: what if referees’ biases are correct, and generally guide them towards making correct calls? Most decisions made by referees are made in an instant, they do not have the luxury of taking their time to sit down and ponder whether a foul is committed, but rather need to rely on sharp senses, experience and gut feelings. So when, for example, African-American football players are generally given more yellow and red cards by referees, then that could just be because they commit more fouls. Moreover, since experienced referees have this knowledge internalized, they automatically are more prone to blow the whistle against African-American players than other races; that is not a problem because it increases the possibility of making correct decisions. If, on the other hand, referees are not able to use their biases to guide their decisions, this could cause them to make more bad calls, or put differently, being unbiased could mean paying an epistemic cost in order to make ethically right decisions. Or even more sharply: combating implicit cognitive biases would be sacrificing correct decisions to satisfy political correctness.

There are two main counter-objections to this. One is that the epistemic cost should be paid rather than the ethical cost. Rules of sports apply to individuals performing individual acts. Regardless of race, colour, age or gender, each athlete should be treated as an individual. If referees were to openly take race, gender or other physical traits into consideration while making calls, this would mean rewriting the rules of all the games we
play in an intuitively quite repugnant direction. Just imagine a referee who was an outspoken racist and whose refereeing was heavily based on his racist attitudes. That would not be in accordance with the rules of soccer where foul play should be called with no reference to race, and it would also be against one of the central ideas of Olympism which we stated in the introduction: ‘Every individual must have the possibility of practicing sport, without discrimination of any kind’. Having a referee that makes the occasional bad call is part of the game, having a referee that is partial is not.

Secondly, even if, for the sake of the argument, we were ready to pay the ethical cost, the alleged epistemic advantages of implicit biases are highly controversial, as it is far from evident that biases are in fact a useful tool in guiding such calls. Humans are, in general, terrible at having their decisions regarding individual cases guided by statistics (Kahneman 2011) and it is at best an open empirical question whether sports referees are better than the average person is at that.

**Objection 4: Insufficient predictive power of the tests**

An important empirical objection is that while implicit bias tests can provide reliable information about implicit bias in the general population, they are not sufficiently accurate to predict the behaviour of individuals (Greely 2013; Kang et al. 2011; Kurdi and Banaji 2017; Roberts 2011). Even though implicit biases are known to affect behaviour, the exact causal relationship between implicit bias and acting on that bias is far from well exposed. Thus, the concern that this objection voices is that an inaccurate test may have adverse consequences for the sports community. Too many false positives may, for example, lead
to spending time and effort on a problem which appears greater than it is. On the other hand, too many false negatives may lead to ignoring a real problem.

These concerns are relevant and the relationship between implicit biases, test results and effects on individual behaviour is a complex one which is unlikely to be completely resolved in the near future. However, that does not mean that tests cannot be a useful tool for educating referees.

To recap: almost everyone is implicitly biased, even people working in areas where impartiality is essential. Implicit biases affect behaviour, and explicit biases can be revealed using tests. In addition, there is a problem of bias in refereeing, such as home team and racial bias. The exact relationship between these facts is still unclear, but leaving that work to the experts, we can ask the question: when is an implicit bias test accurate enough? The short answer is that it is sufficiently accurate when it works to improve refereeing without sacrificing something more important. But as we have argued in this section this does not seem to be the case.

5. Conclusion

In this paper, we have presented an argument for the conclusion that sports organisations ought to require referees to undergo implicit bias testing as part of their mandatory training. We have also presented and critically discussed four objections to our argument, all of which we conclude are flawed. There might of course be other arguments against the mandatory use of implicit bias testing, that we have not taken into account in this article. For example, that there exist other kinds of interventions to reveal and minimise biases among referees. One such intervention could be to improve the diversity among members
of the officiating profession. This could, for example, be done if sports organisations made an extra effort (e.g. through affirmative action) to ensure that the group of referees consists of individuals, that represents the background population. Whether an intervention like this is feasibly and morally more acceptable compared to mandatory implicit bias testing is a topic on another article. The idea with this article has been to start a discussion of how we can minimise the reported biases in sports refereeing. Moreover, given that very little, if anything, is presently being done about implicit biases in refereeing, and given the lessons about implicit bias from the criminal justice system combined with the racial and gender bias in sports, the bar is set quite low for how good the tests need to be to improve the status quo.

References


FIFA. 2019. Diversity and Anti-Discrimination at FIFA. Zurich: FIFA.


---

1 For a critical discussion of Price and Wolfers (2010), see Ingraham (2014).

2 For studies showing that becoming aware of an implicit bias towards a minority will reduce the bias in question, see e.g. Pope, Price, and Wolfers (2018). For examples of methods for dealing with implicit bias, see http://confrontingbias.group.shef.ac.uk/resources/.

3 Thanks to an anonymous reviewer for making us aware of this possibility.

4 For the same line of reasoning, though in connection to technologies such as the use of Hawk eye in tennis see e.g. Collins 2008 and 2019.

5 The cost of neuroimaging is anywhere between $250 and $12,000 depending on the scan type see e.g., https://affordablescan.com/blog/brain-mri-cost/

6 Collins (2010), for example, has argued that it is not made clear to the public that the use of Hawk-Eye is not 100% reliable; see also Nlandu 2012 and Torres 2010 for critical discussions on some of the assumptions about effective behind the use of goal-line technologies.