'*Voetbal Hoort niet bij Robots*': Attitudes Regarding the Use of Artificial Intelligence in Refereeing

Natali Helberger and Brahim Zarouali[1]

1. INTRODUCTION

Bernt Hugenholtz is not only a world-renowned expert and scholar in information law, but also a dedicated soccer fan. At the Institute for Information Law (IViR), he combined both his passions. For years, the IViR soccer team – for which Bernt himself played – competed against other soccer teams, mostly from large Amsterdam law firms. The IViR soccer team had potential and was reasonably successful. Some more time spent on the pitch and less on writing papers might have helped them climb even further in the ranks of legal soccer teams in Amsterdam, and perhaps also … some AI?

Sports may well be the area in which AI made its most publicised appearance (that is if you count chess playing as a sport) and Deep Blue's victory over Kasparov in 1997 is often cited as evidence that AI can outperform humans. Since then, the sports sector has been a place for vibrant AI innovation.[2] Perhaps if we had used AI to scout new postdocs and PhDs more actively for their talents in sports, the IViR soccer team could have repeated (or even trumped) the success of Billy Bean, the famous coach of the Oakland Athletics who, with the help of data analytics, turned a B team into a team of winners.[3] Today, there is a range

1. The authors thank Ljubisa Metikos for his excellent research assistance.
2. For a good overview, see 'Artificial Intelligence. Application to the Sports Industry', *pwc.com*, February 2019.
3. This (true) story inspired M. Lewis to write the successful fiction book *Money Ball*, which was later turned into a film with Brad Pitt in the lead role. See also Caitlin Fitzsimmons, 'What the Moneyball story teaches us about investment', *smh.com.au*, 25 April 2017.

of AI-driven tools that help trainers and clubs to scout for talent and identify the right players.[4]

Some AI, or more generally digital technology, could also have come in handy in training our sturdy players. Thanks to the advances in wearables and data-driven performance monitoring and lifestyle tracking, not only sportspersons but essentially anybody who can afford a fitness tracker can use the power of AI and data to get personalised recommendations for the right diet, training and personal development plan. Had the IViR soccer team been equipped with automated fitness trackers, Bernt could have monitored the progress of his players via a handy dashboard from a comfortable chair in his office.[5]

Let us not forget about the fans. It is not difficult to imagine how much it would have boosted the performance of the IViR soccer team had all those non-soccer playing colleagues back in the office used AI and digital innovation to support their team: think of access to enhanced players' statistics, and the ability to follow your favourite player with AI-controlled cameras[6] and receive a personalised livestream of scores.[7] With the advances in automated content production (robot journalists), the results of the game could even have been immediately shared with the university journal *Folia* and the IViR Twitter account. Not to speak of the many fun applications that would have turned even die-hard non-soccer colleagues into avid followers, such as the possibility to use AI to predict and bet on the outcome of games.[8] There is now even the option to boost morale with some fake crowd cheering. To make up for the lack of audiences

4. NBA Global Scout, HomeCourt analysing footage of users of the app playing basketball, *Sportlogiq*, *sportlogiq.com*, accessed 19 June 2021. Edd Gent, 'How AI is helping sports teams scout star players', *NBC News, nbcnews.com*, 13 May 2018.
5. Again, there are many examples. In addition to the popular fitness trackers such as Fitbit, Garmin or Strava, there are, for example, apps such as Home Court (homecourt.ai, accessed 19 June 2021), a system that uses the camera on a mobile device to turn sports training into a live action video game. The Isotechne platform (isotechne.com/isotechne-platform, accessed 19 June 2021) that can use up to twenty cameras to capture players' performance on the soccer pitch and aggregate that information into a report that coaches can use to improve their players. Liverpool Football Club uses SkillCorner's AI systems (Andrew Cohen, 'Liverpool F.C. Partners With SkillCorner's AI Platform to Extract Data From Matches', *sporttechie.com*) to track players and the ball, allowing coaches to spot the teams' weaknesses on the pitch at a glance. Researchers at KU Leuven even collaborated with SciSports to develop an algorithm ('Talent Captured by Tech: Machine Learning on the Pitch', *kuleuven.be*) that can value a football player's total impact on the game. And Ireland's national rugby team uses connected clothing, (statsports.com, accessed 19 June 2021) to collect better and more accurate data.
6. But see also the glitch with the bald referee at an Inverness Caledonian Thistle FC game, See: James Vincent, 'AI camera operator repeatedly confuses bald head for soccer ball during live stream', *The Verge, theverge.com,* 3 November 2020.
7. N. Gillil, 'How Wimbledon is using AI to enhance the fan experience', *eConsultancy, econsultancy.com*, 10 July 2018.
8. For example, Simplebet (simplebet.io, accessed 19 June 2021) launched FanDuel, a betting application that allows users to bet on basketball players (fanduel.com/, accessed 19 June 2021), and the DFL was involved in the development of 'xGoals' that would '[change] how we evaluate Goalscoring Chances', DFL Deutsche Fußball Liga, 'xGoals: changing how we evaluate goalscoring chances, *dfl.de*, 1 April 2019.

during the Covid pandemic, IBM Watson, for example, have trained an AI model to produce fake crowd cheering during live tennis matches.[9]

Then there is the referee.[10] I still do not exclude the possibility that many referees were simply wrong or biased in favour of the teams of the large, wealthy law firms, and that with different referees, some of the results would have turned out very differently for the IViR team. After all, referees making wrong decisions is not an uncommon problem in soccer, a problem that not only impacts the ego and future motivation of the team, but also – at least for professional soccer teams – have very real economic consequences.[11] Not surprisingly, the use of AI and digital technology to assist or even replace human referees is an area of intense experimentation. Examples range from widely used decision support systems like the video assistant referee (VAR) to fully automated robot umpires, like Trackman in basketball.[12] And also for soccer, some innovation experts predict that by 2030, Premier League games will be officiated by robot referees.[13]

And what if the IViR team had finally decided that soccer was just not its thing? Well, it could have used AI to help create a new, custom-made sport that was better suited to the talents and capabilities of IViR researchers. It would not have been the first time. After all, a sport called Speedgate – a mix of rugby, field hockey and soccer – was invented by a data-driven computer program.[14]

As this short introduction has demonstrated, the use of AI in sports affords a host of interesting opportunities, as well as research questions. As this chapter is a tribute to Bernt, my colleague and mentor and the long-time director of IViR, and because Bernt is one of the fairest academics I know, we focus in this chapter on one particular manifestation of AI in sports, namely the robot referee, and how likely it is that in the not so distant future, Ajax fans will have to get used to a robot referee on the pitch.

In addition to the fact that this is a question that I am fairly certain Bernt has a personal interest in, the case of the robot referee is also an interesting case to gain a deeper understanding of why certain forms of automated decision making (ADM) are or are not societally acceptable.[15] ADM systems are replacing human decision-makers in a whole range of areas, from governments and courtrooms,

9. Jen Booton, 'Empty Stadium? There's an AI for That. How the USTA and IBM Are Keeping the U.S. Open Rocking', *Sport Techie*, *sporttechie.com*, 9 March 2020.
10. Jeff Bakalar, 'Meet the AI that could replace your baseball umpire', *CNET*, *cnet.com*, 13 August 2019.
11. Cedric Gottschalk, Stefan Tewes and Benjamin Niestroj, 'The Innovation of Refereeing in Football Through AI', 6 *International Journal of Innovation and Economic Development* (2020).
12. Jeff Bakalar, 'Meet the AI that could replace your baseball umpire', *CNET*, *cnet.com*, 13 August 2019.
13. Marc Andrada, 'Could Robots Ever Be Referees?', *Sport One*, *sport.one*, 29 October 2018.
14. Joe Lemire, 'The times are changing, The Future of Baseball Has Finally Arrived: Robot Umpires', *Sport Techie*, *sporttechie.com*, 10 July 2019.
15. In this sense see also: Meg Jones and Karen Levy, 'Sporting Chances: Robot Referees and the Automation of Enforcement', *We Robot 2017*, ssrn.com, 29 November 2018.

to media, politics and sports.[16] The ongoing integration of ADM has triggered an intense debate among academics, policymakers and civil society about the opportunities and risks that ADM poses, as well as the conditions under which it is acceptable. For example, applications of ADM in the justice sector already range from decision support systems (such as automated text analysis and the jurisprudence robot of the Dutch public prosecutor) to systems that complement or even replace core functions of the police and judiciary.[17] Although the first research findings seem to suggest that many citizens acknowledge that ADM in justice can help to make fairer decisions, the acceptability of ADM is limited due to the lack of the 'human element'.[18] Interestingly, sport, not unlike the law, is a rule-based activity and one that is still believed to be an inherently human activity.[19] And like law, sport is an area that is characterised by public values such as fairness, justice, dignity and solidarity.[20] Through studying users' attitudes to automated referees in sports, we can refine our understanding of the human factor in decision making, and what the prospects are that in the not so distant future we will be living with fully autonomous AI referees and perhaps also judges.

2. THE ROBOT REFEREE AND OTHER DIGITAL APPLICATIONS

AI, digital technology and even robot referees are no strangers to sports as we know them. Probably the best-known example is the Hawk-Eye electronic line-judge system that allows for the automated tracking of a ball and is used in, for example, tennis, cricket, rugby, volleyball and soccer. The system was used during, for instance, the 2013/14 Premier League season as goal line technology to instantly inform the referee whether a goal had been scored. The then chief executive of the Premier League, Richard Scudamore, welcomed the move towards digital technology as a no-brainer: 'Football is fundamentally a simple game; whichever side scores most goals wins. So, when one is scored, or indeed not

16. AlgorithmWatch and Bertelsmann Stiftung, *Automating Society Report 2020* (2020.automating society.algorithmwatch.org, 2020).
17. European Commission, *Study on the Use of Innovative Technologies in the Justice Field* (Brussels, 2020).
18. Natali Helberger, Theo Araujo and Claes De Vreese, 'Who is the fairest of them all? Public attitudes and expectations regarding automated decision-making', 39 *Computer Law & Security Review* (2020); Glikson and Woolley, 'Human Trust in Artificial Intelligence: Review of Empirical Research', 14 *Academy of Management Annals* (2020); Logg, Minson and Moore, 'Algorithm appreciation: People prefer algorithmic to human judgment', 151 *Organizational Behavior and Human Decision Processes* (2019), 90–103, cited 83 times.
19. Though perhaps not the same for all sports. A study into the performance of Formula One drivers found that the contribution of the human driver to the success of a team has declined over time since at least 1980, going from about 30% driver in early 1980s to about 10% driver today. The rest of the performance stems from the car and the technical team. Andrew, Smith, Sabel and Jones, 'Formula for success: Multilevel modelling of Formula One Driver and Constructor Performance', 12 *Journal of Quantitative Analysis in Sports* (2016).
20. Migle Laukyte, 'Disruptive Technologies and the Sport Ecosystem: A Few Ethical Questions', 5(24) *Philosophies* (2020), 1–13.

scored, and we have the ability through technology to definitively know whether the ball crossed the line we should absolutely use it.'[21] Later, after some initial reluctance, also the German Bundesliga (Germany's premier football league) approved the new goal line technology.[22] And in baseball, a related technology – the Pitch f/x camera technology – has made it into the major league stadia to track the speed and location of the ball and determine strike calls and pitches.[23] The data generated can support umpires and sport reporters, but is also being used to analyse games and perform all kinds of data-driven evaluation of the players and, interestingly, also of the performance of umpires themselves.[24]

The VAR system is another example of the use of digital technology and AI in refereeing. It was introduced in the 2018 Soccer World Cup and is now used in various club competitions. VAR originally emerged as part of a project conducted by the Royal Netherlands Football Association (KNVB) under the heading 'Refereeing 2.0'.[25] The VAR system, in cooperation with three human VAR assistants, is used to replay and double-check the referee's decisions and can, in the case of clear and obvious errors, lead to a decision being revisited or overturned. And if today players are clustered around the referee making a TV sign with their fingers, it is because they want a situation to be revised with the help of the VAR. VAR technology is also one of the first digital technologies to have made it officially into the 'rules of the game'. According to the VAR protocol of the International Football Associate Board (IFAB), '[a] video assistant referee (VAR) is a match official, with independent access to match footage, who may assist the referee only in the event of a "clear and obvious error" or "serious missed incident" in relation to: a. Goal/no goal, b. Penalty/no penalty, c. Direct red card (not second yellow card/caution), d. Mistaken identity (when the referee cautions or sends off the wrong player of the offending team).' Ultimately, however the final decision is always taken by the human referee.'[26] By the way, players who excessively call for a VAR 'review' can be given a caution.[27]

In addition to real-time judgement or evaluation based on past data, AI and digital technology can also be used in a more predictive way. The xGoals algorithm, for example, can be used to determine the probability of a player scoring a goal in a particular situation.[28] According to the DFL (German Football League), the technology compares a shot against more than 40,000 similar efforts stored in the Bundesliga's official database and can be used to assess the

21. Cited in: Owen Gibson, 'Premier League clubs choose Hawk-Eye to provide new goalline technology', *The Guardian*, 11 April 2013.
22. 'Bundesliga approves Hawk-Eye goal-line technology for new season', *Carley Observer, carlyleobserver.*com, 4 December 2014.
23. Mike Fast, 'What the Heck is PITCHf/x?', *The Hardball Times Baseball Annual* (2010).
24. N. DiMeo, 'Pitch f/x, the new technology that will change baseball analysis forever', *slate.com*, 15 August 2007.
25. João Medeiros, 'The inside story of how FIFA's controversial VAR system was born', *Wired UK, wired.com*, 23 June 2018.
26. IFAB, 'Video Assistant Referee (VAR) protocol', *www.theifab.com*.
27. IFAB, 'Fouls and Misconduct', *www.theifab.com*.
28. DFL Deutsche Fußball Liga, 'xGoals: changing how we evaluate goalscoring chances, *dfl.de*, 1 April 2019.

performance of the players and the goalkeeper,[29] as well as to examine whether the actual score in a match is a 'fair reflection of what took place on the pitch'.[30]

Technology has also proven capable of autonomously performing refereeing tasks during matches, without any human input. An example of this is the use of the Trackman automated ball-strike system in the Atlantic League's All-Star baseball games. This automated umpire sends its decisions regarding ball and strike calls to a human umpire on the field. This human umpire does not check or change the decision, but merely relays it to the players.[31] The use of Trackman was only experimental, however. A full-scale permanent implementation of autonomous AI is yet to be seen.

3. BUT … DO WE WANT IT?

There are many good reasons why we may want to support the use of AI and digital technology and maybe one day even want to replace referees and umpires with AI-driven refereeing systems. One important reason is to increase the accuracy and fairness of decision making. According to Lukas Brud, secretary of the IFAB: 'If we had mentioned the idea of introducing video referees in 2010, people would say we were crazy, but now they saw it as an opportunity to help referees and to achieve fairer outcomes in a match.'[32] An analysis by sport scientists at KU Leuven of more than 800 soccer matches in more than twenty countries found that thanks to the use of VAR, the total accuracy of refereeing decisions had increased from 93% to nearly 99%,[33] especially in high-paced games like soccer, tennis and baseball. Gottschalk, Thewes and Nistroy point to the fact that in one soccer game, a referee has to make about 200 distinct decisions, each within a fraction of a second.[34]

Human referees are, moreover, subject to human biases and imperfections. According to Lindholm, human umpires make wrong calls 15% of the time,[35] and Jones and Levy's literature review revealed that factors that create biases in human referees can range from the home crowd's noise and the reputation of a player, to a team's origin and the player's/referee's racial differences.[36]

---

29. *Ibid.*
30. *Ibid.*
31. Zack Willis, 'How Did Baseball's Robot Umpires do, and are They the Future for MLB?', *www.sportscasting.com*, 10 August 2019.
32. Cited in https://www.wired.co.uk/article/var-football-world-cup.
33. Helsen, Gilis and Weston, 'Errors in judging "offside" in football: Test of the optical error versus the perceptual flash-lag hypothesis', 24 *Journal of Sports Sciences* (2006), 512–528.
34. Gottschalk, Tewes and Niestroj, *supra* note 11, 35–54.
35. S. Lindholm, 'How Well Do Umpires Call Balls and Strikes?' *SBNation*, *beyondtheboxscore.com*, 27 January 2014.
36. Jones and Levy, 'Sporting Chances: Robot Referees and the Automation of Enforcement' *We Robot 2017*, ssrn.com, 29 November 2018. Erikstad and Johansen, 'Referee Bias in Professional Football: Favoritism Toward Successful Teams in Potential Penalty Situations', 2 *Frontiers in Sports and Active Living* (2020), https:/www.frontiersin.org/articles/10.3389/fspor.2020.00019/full#B18.

In addition, Gottschalk et al. refer to the immense psychological pressure on referees and argue that the transition to robot judges can reduce economic harm, psychological pressure and the risk of manipulation.[37] According to Lukas Brud: 'We knew we had to protect referees from making mistakes that everyone can see immediately.'[38] Having said that, AI driven technology can be used not only to assist referees but also to evaluate them, and thereby potentially increase psychological pressure. In an article reviewing the first use of VAR technology during the 2018 World Cup, the journalist ended with: 'Another thing to consider would be to stop those such as Cáceres from officiating top-level matches again. The 44-year-old's use of VAR on Monday exposed him as a weak-minded official, turning something designed to make his job easier into a tool for chaos.'[39]

The general perception that the introduction of AI technology in decision making can result in fairer outcomes is echoed by a growing body of research about user perceptions of ADM. Studies show that under certain conditions, people prefer advice from an ADM to advice from humans, a phenomenon that Logg et al. call 'algorithmic appreciation'.[40] Also in the context of algorithmic decision making in the media,[41] people tended to favour an algorithmic decision to that of a human decision maker under certain conditions. This is not to say that ADM cannot be afflicted by all sorts of biases.[42] Shin shows how, also because of a growing awareness of bias and other failures in ADM, fairness, accountability, transparency and explainability (FATE) principles are an important parameter for engendering users' trust in ADM, because users rely on FATE as heuristic tools to assess the trustworthiness of algorithms.[43]

One important factor that contributes to algorithmic appreciation in the justice sector is that people can find the decisions of an AI judge fairer than those of a human judge. Having said this, and interestingly, even in situations in which AI is considered to be able to make the fairer decisions, this does not automatically mean that people are willing to accept ADM. Whether they are willing to do so also depends on the extent to which other factors are present, such as respect for human dignity, the ability to express and understand emotions, and the human

37. Gottschalk, Tewes and Niestroj, *supra* note 11, 48.
38. João Medeiros, 'The inside story of how FIFA's controversial VAR system was born', *Wired UK, wired.com,* 23 June 2018.
39. Sachin Nakrani, 'VAR: the World Cup verdict so far – some success but more clarity needed', *The Guardian,* 26 May 2018.
40. Jennifer M. Logg, Julia A. Minsona and Don A. Moore, 'Algorithm appreciation: People prefer algorithmic to human judgment', 151 *Organizational Behavior and Human Decision Processes* (2019), 90–103.
41. Thurman, Moeller, Helberger and Trilling, 'My Friends, Editors, Algorithms, and I', 7 *Digital Journalism* (2019).
42. Lepri, Oliver, Letouzé, Pentland and Vinck, 'Fair, Transparent, and Accountable Algorithmic Decision-making Processes, The Premise, the Proposed Solutions, and the Open Challenges', 31 *Philosophy & Technology* (2018).
43. Donghee Shin, 'User Perceptions of Algorithmic Decisions in the Personalized AI System: Perceptual Evaluation of Fairness, Accountability, Transparency, and Explainability', 64 *Journal of Broadcasting & Electronic Media* (2020).

touch of the decision maker.[44] According to Lee, much also depends on the characteristics of the task in question and whether it requires more 'human' or more 'mechanical' skills. For tasks that mainly involve mechanical skills, users can find an ADM system just as fair as a human decision maker, even though the reasons behind people's perceptions differed (for human-made decisions, fairness and trust are attributed to managerial skills, and for ADM decisions to reliability and the lack of bias).[45] However with tasks that require human skills, people valued ADM less positively because they lacked human intuition, relied solely on quantifiable metrics, and were unable to evaluate social interaction or handle exceptions, or were demeaning and dehumanising when used to judge people.[46]

The human factor is also a central factor in sports. Laukyte argues that sport is done by humans for humans, and that '[t]aking a human judge out of the loop means, to a certain extent, dehumanizing sport and inverting the praxis according to which it is humans who usually evaluate technology'.[47] In a similar vein, Jones and Levy, while acknowledging that an AI umpire can be the fairer decision maker, point to the importance of the human sociological context. They argue that the preservation of beneficial unpredictability, or what they call the 'sporting chance', is a reason for not appreciating automated refereeing. As their argument goes, even if an algorithmic system like a VAR were in a better position to make accurate and fair decisions, it is important to not underestimate the value of imperfect enforcement. Their analysis of four sports (golf, tennis, basketball and soccer) revealed that digital technologies and AI are used in different sports to enforce rules or aid umpires, and in so doing can indeed result in better, fairer and more accurate decisions, but that they also reduce the fun of watching the drama when tennis stars like Connors or McEnroe smash their rackets and try to intimidate the umpire. More concretely, they elaborate on six such elements, namely drama, adversity, custom, integrity, humanity and dignity.

Concerns about what the introduction of AI technology would do to a sport as such (i.e. the aspect of custom) were also raised in the context of the introduction of the VAR system, and more generally, probably any technological innovation.[48]

So while AI in sports refereeing can have clear benefits as well as drawbacks and less desirable implications, whether to use or continue to use AI-driven technology is, ultimately, a decision made by humans. And in this context the

---

44. Helberger, Araujo and De Vreese, *supra* note 18.
45. Min Kyung Lee, 'Understanding Perception of Algorithmic Decisions: Fairness, Trust, and Emotion in Response to Algorithmic Management', 5(1) *Big Data & Society* (2018), 11, even though there is growing evidence that people are starting to understand the factual wrongness of the claim that AI decisions are characterised by a lack of bias, see Bruno Lepri, Nuria Oliver, Emmanuel Letouzé, Alex Pentland and Patrick Vinck, 'Fair, Transparent, and Accountable Algorithmic Decision-making Processes', 31 *Philosophy & Technology* (2018), 611–627.
46. Lee, *supra* note 45, 12.
47. *Ibid*., 10.
48. Tom Lamont, 'He shoots, he scores – or does he? How VAR changed football for ever', *The Guardian*, 23 January 2021.

audience has undoubtedly an important role: if spectators do not like it, it is unlikely that we will see a further evolution of AI referees in sports. Therefore, and inspired by the theoretical assumptions made by Jones and Levy (and others), we investigated attitudes towards and the acceptability of AI referees in sports among a representative sample of Dutch citizens. In the following, we report some key insights from that research. In so doing, we also contribute to a growing body of research into human attitudes and understandings of automation and rules, and what exactly that human factor is.

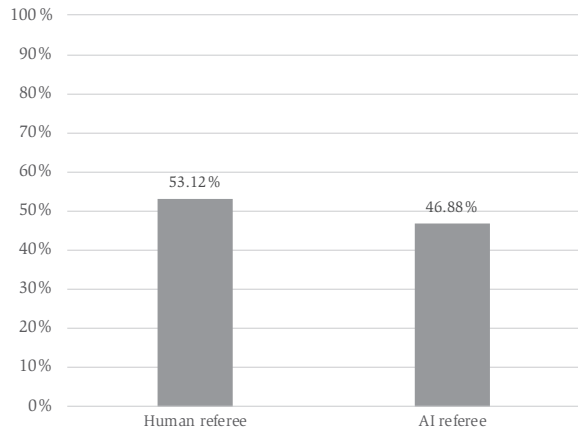
4. REPORTING FROM A SURVEY

To explore what people think about AI referees, we conducted a survey among a representative sample of the Dutch population. Fieldwork was carried out by research company IPSOS in April–May 2021. The total sample size was 1,009. The respondents had a mean age of 48.2 years (*SD* = 16.7 years) and half of them (50%) were women. Around 17% had a lower education level, 40% a moderate one and 43% a higher one. Prior to the start of the survey, respondents were given a brief introductory text explaining the concept of AI referees. They were informed about what AI referees are, how they are being used (e.g. to see whether a ball crossed the goal line in football) and how they sometimes replace real referees (since AI is becoming increasingly smarter). The introductory text familiarised the respondents with the topic, ensuring that they could answer all questions in a more informed way.

4.1 Human vs. AI Referee: Which One is Fairer?

The first set of questions concerned whether people thought a human referee or an AI referee would make fairer decisions. To trigger people to consider the

Figure 1: Perceived fairness of human referees vs. AI referees



differences between a human and an AI referee, we asked a very binary question and did not offer the option of a hybrid relationship in which the human referee cooperates with an AI referee. As illustrated in Figure 1, respondents were almost equally divided between those who considered the human referee fairer (53.12%) and those who considered the AI referee fairer (46.88%).

4.2 Reasons for Considering a Human Referee or an AI Referee Fairer

We then asked the respondents to briefly explain their answer ('Could you briefly explain why you think that?'). We manually inspected all the answers, and the main arguments of those who considered the AI referee fairer could be roughly divided into five categories: 1) absence of biases (230 responses), 2) accuracy of decisions (93 responses), 3) ability to see more (72 responses), 4) consistency in decisions (7 responses) and 5) AI is the future (3 responses). The main arguments of those who considered the human referee fairer can be clustered around very different answer categories, namely 1) human ability to judge context (181 responses), 2) human ability give emotions a role in decisions (107 responses), 3) human ability to make exceptions (24 responses) and 4) human ability to provide communication (14 responses). We now provide some insightful answers ('quotes') to discuss these findings.

Regarding the arguments in favour of AI, the alleged non-bias of AI was the most cited reason (230 responses). A significant share of respondents said they believed AI referees to be unbiased. Some cited as reason the lack of emotions in AI: 'Refereeing is about rules. And only a computer can enforce that perfectly. Because a computer has no feelings, no gut feeling, and is purely objective' (respondent 84). Others mentioned the lack of prejudices and personal preferences, or as one respondent put it: 'AI can't be a fan of a specific club' (respondent 558).

Interestingly, the ability to have emotions (107 responses) and consider the human context (181 responses) were also often cited as reasons why human referees are fairer. As one respondent observed, fairness is more than accuracy and following the rules: 'With the help of AI, things can be perceived much more accurately, but honestly, there is more than just accuracy and following the rules; sports is related to emotion, weighing a judgment in certain context' (respondent 644).

When it comes to accuracy, several respondents said that AI is more precise, more objective, lacks emotions and is not prone to making human mistakes: 'The human referee makes mistakes and makes decisions based on feelings, which is almost impossible with an AI referee' (respondent 130). Often, however, it was human imperfections rather than perfect AI that caused respondents to consider an AI referee fairer: 'A referee is only human and can therefore make human errors. A robot can also get stuck or something, but in principle, it is adjusted to do the right thing and that is more fair' (respondent 323).

Even though many respondents conceded that an AI referee may make accurate decisions, they also found that it is less able to judge the broader context: 'A human referee looks not only at the moment, but also at the match as

a whole. Sometimes you have to be very strict, and sometimes you don't have to' (respondent 325). In addition to context, the ability to take decisions 'in the spirit of the game' was repeatedly referred to and a factor that could even outweigh the occasional mistake: 'A human referee makes decisions "in the spirit of the match". Any errors that occur are part of the sport' (respondent 218). Sometimes, however, decisions are not clear, and here it is the ability to engage in an argument and communicate with a judge that was a reason for a number of respondents to consider a human fairer: 'The human referee always remains flexible, but a robot acts in black or white. You can't argue with this either. Machines can also make mistakes. It doesn't get any more fun' (respondent 143).

Interestingly, some respondents clearly nuanced and differentiated between situations in which an AI or a human referee is fairer: 'In sports too, just like in real life, there are nuances. A referee must also be able to feel a match, so to speak. It is more than strictly following the rules. Rules are there to keep the game fair and orderly. That is a means and not an end in itself. So also in sports the following applies: it is not science but reasonable knowledge. AI can help the human referee though' (respondent 763). Accordingly, some respondents also argued that it is the combination of AI and human that can result in fairer decision making: 'When it comes to behaviour – for example, a foul – I prefer the human referee because the "spirit of the match" is always involved. In decisions with a clear boundary – for instance, ball in or out – the AI will always be fair' (respondent 46).

4.3 Acceptance and Trust of Human/AI Referee's Decisions

After the open question, we asked respondents about the extent to which they accept and trust decisions made by human vs. AI referees. When it comes to acceptance (see Figure 2), interestingly, the willingness to accept the decision of an AI referee was rather high. That is, around 62% of the respondents were at least rather likely to accept an AI referee's decision vs. only 14% who were not inclined to accept the decision of an AI judge; around 24% of the respondents were neutral. When asked whether people were willing to accept a human decision supported by AI, the number of respondents who answered 'very likely' was much higher (in total, around 77% indicated their willingness to accept such a hybrid decision).

In the context of trust, we compared a human referee with an AI referee. Respondents were asked to what extent they agreed with the following statement: 'I trust the AI referee OR human referee to be right' (see Figure 3). The answers 'rather agree' and 'strongly agree' are in favour of either the human or the AI referee. However, we also observe that in terms of total agreement (i.e. if we look at all three 'agree' answers), people were almost equally inclined to trust the decision of an AI referee (64%) vs. a human referee (71%).

Figure 2: Acceptance of decisions made by an AI referee vs. a human referee + AI

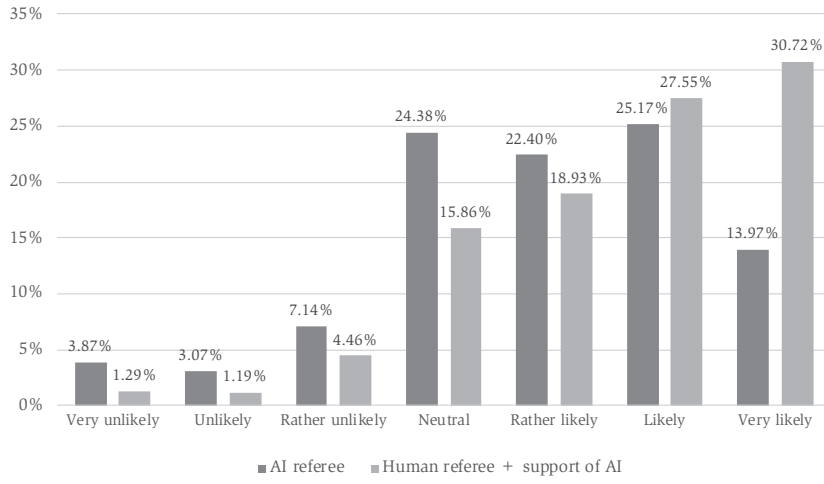
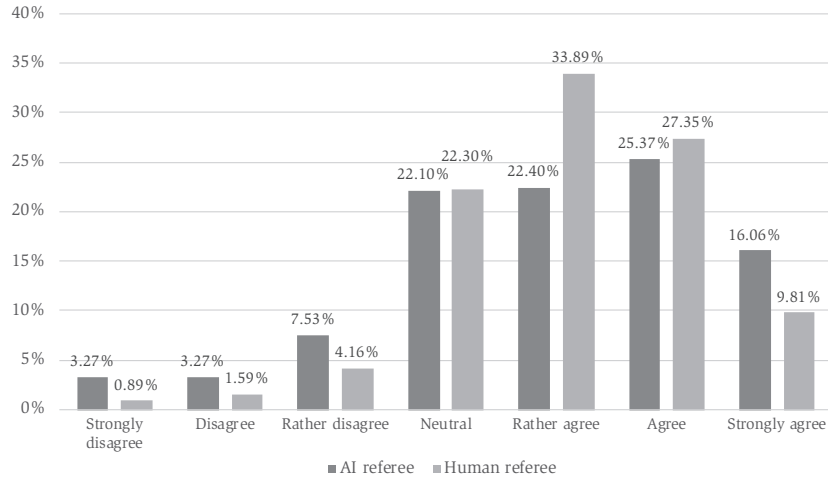


Figure 3: Trust in an AI referee vs. a human referee



4.4 Human Agency

We then asked about human agency and the 'human in the loop' that is also mentioned in both the literature and regulations as an important normative requirement. In this section, when we mention 'total agreement', we refer to the answers 'rather agree', 'agree' and 'strongly agree' combined. The results indicated that, also for the case of using AI in sports, there was broad agreement

among our respondents that, ultimately, the human should be in control (32% of respondents strongly agreed with this proposition, and there was 83% total agreement). Also, only 6% strongly agreed that an AI referee should always have the last word (and 36% total agreement). In addition, 31% of the sample strongly agreed that the human should have the last word (74% total agreement). Very interestingly, however, there was rather broader agreement that also the human referee should be controlled through an AI system ('AI in the loop') (19% strongly agree, and 63% in total agreed).

4.5 Reasons for Preferring an AI Referee

We then looked into the six elements of sporting chance developed by Jones and Levy as reasons why people might prefer a human referee over an AI referee even if the latter is considered to be more accurate and/or fairer or less biased.[49] These elements are drama, adversity, custom, integrity, humanity and dignity. In the light of this, we presented the respondents with 10 statements that tapped into their main reasons for preferring a human or an AI referee. These statements were measured on a 7-point Likert scale, ranging from 1 (= strongly disagree) to 7 (= strongly agree). The statements are presented in Table 1.

Table 1: Overview of all the statements and their agreement level

Statement	Strongly disagree	Disagree	Rather disagree	Neutral	Rather agree	Agree	Strongly agree
AI referees will make sports competitions fairer	4.7%	2.2%	7.8%	28.1%	24.0%	20.1%	13.2%
AI referees are less biased	3.3%	1.5%	4.6%	20.9%	22.0%	21.1%	26.7%
AI referees are more accurate	3.6%	2.7%	6.0%	27.8%	22.8%	22.4%	14.9%
AI referees are more reliable	4.6%	3.2%	8.7%	32.5%	22.4%	18.9%	9.7%
AI referees will make sports matches more exciting	12.6%	10.0%	16.4%	36.1%	14.0%	6.8%	4.2%
AI referees will make sports matches more fun to watch	13.4%	11.7%	15.2%	34.0%	14.7%	7.1%	4.0%
Disagreeing with the referee and opposing his/her decision is part of the game	7.6%	7.3%	8.5%	24.9%	22.9%	15.8%	13.0%
Sports competitions are no longer the same with AI referees	4.1%	5.7%	6.8%	25.2%	21.2%	17.0%	20.1%
Sport is 'human', and so should be the referees	1.6%	2.2%	4.3%	19.7%	22.2%	20.3%	29.7%
Human athletes should not be judged by an AI	8.82%	11.40%	14.47%	30.23%	15.06%	9.91%	10.11%

49. Jones and Levy, *supra* note 15, 3.

Without going into too much detail, a couple of interesting trends can be drawn from Table 1. On the one hand, a considerable share of the respondents said that AI referees are less biased, more accurate and more reliable, and that thus AI referees could make sports competitions fairer. Having said this, the respondents did not agree that much with the idea that AI referees will make sports more exciting or more fun to watch. Many respondents also said that the human element is still very important. For instance, 'disagreeing with a referee and opposing his/her decision' is something that people engage in very often. Also, a significant share of the respondents said that sports competitions would not be the same with the involvement of an AI referee, so the human touch remains important here. When it comes to the general idea that 'human athletes should not be judged by AI', there is no real consensus: the respondents were very divided about this, which means that people might be open to AI-based sports judgments, but this will very much depend on the implementation of the AI, the context in which it is used (e.g. for what kind of decisions) and many other things.

5. DISCUSSION

A quote from a recent sports blog gives a good impression of the growing circle of enthusiastic followers who see a real future in sports for the AI referee:

> In the future Artificial Intelligence will be able to make much better and more accurate decisions than the human referee ever can. Don't be surprised if robots will take over in the future with accurate AI driven decisions. Does it mean, there is no role left for the referee? No! I believe he will stay around, only in a different role. The role of the referee in the future will be one that is mainly related to human behaviour, emotions and conduct (which AI is still not fully capable of), unless we decide differently of course.[50]

Judging by our results, human referees are here to stay, and they will do more than merely lend emotional support. And judging by our respondents' answers, it is very unlikely that we will decide differently at any time soon, at least not in the Netherlands. Having said that, our respondents broadly agreed that AI can and should help human referees to make fairer and better decisions. Our survey thus revealed a rather high level of trust in and willingness to accept AI-assisted decisions.

It is indeed worth pointing out that our respondents valued different qualities of an AI vs. a human referee and that the clear majority favoured a hybrid model of AI-assisted decision making. This is a model in which the human remains in control. However, a significant share of respondents also argued that it should be the other round, too; namely AI should be used to control human referees, very much in the way that VAR is used. Indeed, as one respondent explained: 'I have been a referee myself, for more than 35 years. Having only

50. A. Fetters, 'How Instant Replays Changed Professional Tennis', *www.theatlantic.com*, 7 September 2012.

a robot as referee is not a good idea because you are then missing the element of human perception and feeling for the game. VAR technology is becoming better to complement selected moments for the referee. VAR is great for offside' (respondent 481).[51]

Curiously, we also saw something that can probably best be described as grudging tech resignation: even though the majority of respondents agreed that AI clearly has a role to play in judging sporting competitions, they also largely agreed that using AI in sports will change sports, and not for the better. Games will be less fun to watch, not in the least because disagreeing with the referee is part of what makes sports fun, but also because sports are human. Or as one of the respondents put it so succinctly: *Voetbal hoort niet bij robots*.

51. See also Gottschalk, Tewes and Niestroj, *supra* note 11, 49, also pointing to the need to differentiate between different kinds of decisions.