

Rules, Standards, and the Video Assistant Referee in Football

*Jan Zglinski**

published in: *Sport, Ethics and Philosophy* (2020)[†]

Abstract: Introduced with the hope of reducing refereeing errors and increasing “football justice”, the Video Assistant Referee (VAR) has attracted much criticism from players and spectators alike. Drawing on evidence from domestic and international competitions from the past three years, this article investigates the problems that have become apparent with the system. It argues that the success of technological aids like the VAR depends on the nature of the norms on which they adjudicate. Just like legal norms, football laws can be divided into rules and standards. While the VAR has the ability to make a substantial contribution to enforcing the former, its added value for policing the latter is more limited.

Introduction

The Video Assistant Referee (VAR) is arguably the most significant development in football officiating since the introduction of yellow and red cards in the 1970s. After initial pushback from football’s central governance body FIFA, it was adopted in 2018, with some national leagues already using the system the preceding year. Its design resembles the various instant replay mechanisms that have become common in many sports, such as tennis, basketball, cricket, and American football. The VAR watches the game from a dedicated video operating room, with access to live footage of the on-pitch action. Its task is to alert the head referee of “clear and obvious errors” and “serious missed incidents” relating to four reviewable calls – goal/no goal, penalty kick decisions, red cards, and cases of mistaken identity – which are deemed to represent the most significant refereeing decisions in football.¹

* London School of Economics. Email: j.zglinski@lse.ac.uk.

† Available online at www.tandfonline.com/doi/abs/10.1080/17511321.2020.1857823?src=&journalCode=rsep20.

The introduction of the VAR and, a couple of years earlier, that of the Goal Line Technology (GLT), triggered a debate as to the promise and perils of technological aids in football. Advocates argued that the reforms would reduce the number of officiating errors and, in this way, increase “football justice” (Simòn 2019; Ryall 2012). Critics warned against exceedingly technologizing the sport. Nlandu (2012) suggested that the reforms were motivated by the erroneous belief that refereeing decisions are of central importance for the outcome of a football game and rested on the equally wrong assumption that some match situations affect the result more than others. Collins (2010) claimed, based on an analysis of systems set up in other sports, that the technological devices would create a “false transparency” as they would have significant error margins that would not be displayed to the public.

This article does not seek to revisit the question as to whether technological aids should be employed in football but will provide a preliminary assessment of the VAR’s functioning, based on the experience with the system in national and international competitions over the past three years. What contribution has the VAR made to policing infractions of football norms? In which areas has it been effective, in which ones has it not, and why? The thesis put forward is that the technology has had a positive impact on officiating in relation to rule-like football laws, while having a much more limited success when it comes to standard-like offences. In what follows, I shall first explain what is meant by these terms and show that football norms, just as legal norms, can be divided into rules and standards. I will then go on to argue that their respective properties have repercussions for the incidents that can be reviewed by the video referee and the threshold for VAR intervention, and address some of the recently made proposals to reform the system.

1. Rules and Standards

When talking about legal directives, lawyers distinguish between rules and standards. Both types follow an “If X, then Y” model (Ehrlich and Posner 1974; Schlag 1985), but the respective X – in some cases, also the Y – differ. Rules are

structured as follows: if a specific fact occurs (or a set of facts), this triggers a certain pre-defined legal response. The example most commonly used in legal scholarship is speed limits, such as “it is prohibited to drive faster than 30 miles per hour” or “those driving fast than 30 miles per hour will receive a £100 fine”. To find out whether a motorist has violated this rule one just needs to answer the question as to whether the triggering fact has occurred: does the car move at a speed greater than 30 miles per hour? If so, the legal consequence takes effect (prohibition or fine).

Standards are more complex in nature. Instead of laying down factual criteria that will, once met, automatically prompt a certain response, they require the decision-maker to apply some background principles, which encapsulate the rationale behind the norm, to the case at hand (Sullivan 1992; Korobkin 2000). The law of negligence is a good example. It is, in many countries, defined as not acting as a “reasonable person” would. When deciding whether a certain conduct was or was not negligent, there is no one simple fact that a judge can rely on. They must consider whether the precautions taken, the foreseeability of the conduct, the harm done, and similar considerations warrant punishing the defendant.

The difference between rules and standards can be made sense of as giving a norm content *ex ante* or *ex post* (Kaplow 1992). While the normative substance of a rule is clearly set out in advance, i.e. before any individual conduct has even taken place, standards are normatively under-determinate and must be concretized at the stage of application, when the legality of a specific act is assessed. From an institutional perspective, this means that rule-like norms are fully articulated by the law-maker, which conclusively decides what type of conduct is prohibited. Standard-like norms, by way of contrast, are only “sketched” by the law-maker and primarily shaped by the institution charged with enforcing them (usually a court or agency). Moreover, the choice between rules and standards affects what it is that the enforcing authority has to do. When it comes to rules, the only thing which is required is a factual assessment or, more precisely, an assessment of whether the

triggering fact has occurred. Standard-based decisions likewise require some fact-finding but, in addition, hinge on an evaluative judgment that will incorporate and reconcile the principles behind the norm.

Note that the distinction between the two types is not always clear cut (Schlag 1985; Schauer 2003). Rules often have exceptions that make it necessary to consider additional facts going beyond the narrow triggering criteria (e.g. it is illegal to punch someone unless for self-defence) or qualifications which are formulated in a standard-like fashion (e.g. all foreign marriages are recognized except those which are contrary to “public order”).² Standards, by way of contrast, can contain elements of rules, such as specific tests or guidelines, or can acquire a rule-like nature over time once a sufficient number of precedents clarifies their meaning. Therefore, the two should be thought of as the extremes of a sliding scale rather than in strictly dichotomous terms.³

One misunderstanding should, however, be avoided. The distinction is not to be confused with the one drawn by Russell (1999) between rules and principles. Inspired by Dworkin (1977), Russell has argued that sports are not just governed by formal norms, i.e. those explicitly laid down in the official handbooks, but also by unwritten principles such as fair play and sportsmanship. The rules/standards-distinction, at least in theory, applies to all football norms, even if in the following the focus will be on those of the formal kind. *Vice versa*, the role of principles is more pronounced in relation to standard-based calls, where principles can help the referee concretise the given norm.

Both rules and standards have their advantages and disadvantages. Rules communicate clearly that a certain behaviour is (or is not) allowed and, thus, establish a strong incentive (or deterrent) for engaging in it. As a consequence, they reduce the number of legal disputes and, where these do arise, are comparatively easy to apply and produce predictable outcomes. However, rules also have drawbacks. They can be hard to create in the first place as the law-maker must agree

in advance on all the considerations that are relevant for assessing a certain scenario. More importantly, they tend to cause problems of over- and under-inclusion (Schauer 1991). The clear-cut consequences they stipulate can, in borderline cases, appear unfair, even at odds with the principles motivating them. Standards prevent this issue by letting the institution enforcing them decide on a case-by-case basis, and upon consideration of all pertinent concerns, what behaviour is acceptable. Because of that, however, the decision also becomes less predictable and more challenging to make.

The efficiency of each type of norm depends on the given context (Korobkin 2000). Rules tend to work better in areas which are well understood and marked by similar factual patterns, whereas standards are more appropriate for fields that are factually heterogenous or tend to evolve quickly, for example due to scientific progress. These strengths and weaknesses are often considered by law-makers before formulating a norm, in order to ensure that the most efficient solution is adopted. It is a fascinating question, but one which cannot be addressed here, which type of norm is most appropriate under which conditions for regulating social behaviour in football. Instead, the focus in the following shall be on the implications of the choice, once made, of one over the other on officiating. How does formulating a football norm as a rule or a standard affect the work of referees and video assistants? To understand this issue, we have to take a closer look at the nature of football norms first.

2. The Laws of the Game

Football is governed by the so-called “Laws of the Game”, a collection of norms dating to 1863 which is controlled and annually updated by the International Football Association Board (IFAB). Similar in format to a statute or code, they contain 18 laws that set out authoritatively how football is and is not to be played. This includes organizational matters, such as the pitch size, the location of field markings, and the dimensions and quality of the ball. More importantly for

everyday officiating, they also determine questions relating to scoring and permissible behaviour during a match.

Just like legal norms, the Laws of the Game can be divided into rules and standards.⁴ Perhaps the most famous rule is that “a goal is scored when the whole of the ball passes over the goal line”. The response (goal) is triggered if the pre-defined fact (ball is over the line) occurs. Another well-known and more frequently applied rule is that of (active⁵) offside. It prohibits passes to players who are, with any part of their body with which they could potentially score, nearer to the opponent’s goal line than both the ball and the second-last opponent. Again, the triggering fact (passing to a player who is in an offside position) automatically leads to a specific response (free kick for the defending side).

Other football laws are standards. Early versions of the Laws of the Game prohibited “ungentlemanly behaviour”, an offence which has in the meantime been replaced with that of “unsporting behaviour”. What counts as ungentlemanly or unsporting has never been laid down, even if an indicative list of prohibited conduct has emerged over time,⁶ and is to be determined based on ideas of fairness, sportsmanship, and decency. Another law makes it a red-card offence to engage in “serious foul play”, which is defined as a challenge that “endangers the safety” of an opponent or uses “excessive force or brutality”, broad notions which require normative concretization by the referee. Crucially, most infractions resulting in a penalty kick are formulated as standards. This includes handball offences, which lead to a penalty if a player of the defending side “deliberately” touches the ball in the penalty area with their hand/arm or touches the ball in a way that makes their body “unnaturally bigger”. It also extends to penalties resulting from a foul, which require that the defender kicks, pushes, or tackles the opponent in a manner that is “careless, reckless or using excessive force”.

It is worthwhile noting that this categorization is not static. Over time, many standards have, partly or completely, turned into rules. (The opposite evolution,

from rule to standard, is rarer but can also be observed at times.⁷⁾ This notably happens when guidelines explicitly prohibiting a certain type of conduct are added to and, thus, concretize an otherwise open-textured football law. For instance, in the late 1990s, after a rise in aggressive fouls, the aforementioned general definition of “serious foul play” was complemented with a provision specifically stipulating that studs-up tackles from behind would be punished by sending off. A few years later, an amendment to the Laws of the Game clarified that removing one’s jersey during a goal celebration would, from now on, constitute “unsporting behaviour” and prompt a yellow card. In cases like these, a rule is carved out for a specific behaviour within the broader normative space of a standard.

The choice between formulating a football norm as a rule or a standard has a series of implications for officiating. When it comes to rule-like offences, the main focus is on the occurrence of the given situation or prohibited action. From the refereeing side, this means that a factual assessment is at the heart of the call. If the ball passes the goal line between the goalposts in its entirety, the attacking side has scored a goal. If a player takes off their shirt after scoring, they will receive a yellow card. It does not matter whether the ball has crossed the line merely by a couple of millimetres. Nor does it matter that the jersey was just off for a few seconds or an endearing message was hidden underneath it. The fulfilment of the law’s factual criteria automatically leads to the award or sanction. This might, at times, feel harsh (“it was so close!”) or even unfair (“but he or she did not mean to!”). Alas, this is the nature of rules. Once the facts have been established, the consequence is unavoidable, even if, in atypical or marginal cases, there will be costs of over- or under-inclusion.

The situation is different for standards. Here, a factual assessment is required, too, but it is only the first and, typically, not the most important element of the referee’s decision. Take penalty kicks from a foul. The referee must, to begin with, determine whether there was physical contact between the defender and the opponent: has the former kicked, pushed, or tackled the latter? This is the easy and usually clear part

of the call. Much trickier and more significant is the evaluative side: was this done in a manner that is “careless, reckless or using excessive force”? This requires an analysis of factors such as the strength of the contact, the speed of the action, and the degree to which it has impacted on the opponent’s goal-scoring attempt. There is a continuum ranging from light challenges with a minimal effect on the striker to openly violent tackles that completely foreclose the opportunity to score. Where the line between permissible and punishable action ought to be drawn is not laid down in the Laws of the Game. It is, as sports commentators say, a “judgment call”.

As a consequence, standard-based decisions differ from rule-based calls in two ways. Institutionally, instead of football’s “law-maker”, the IFAB, determining what constitutes permissible and what impermissible behaviour, the decision is delegated to the referee. Substantively, referees have discretion when it comes to shaping the given football law’s normative content. This is not to say that they have an unfettered interpretive freedom; certain calls are outside their scope of discretion. For example, under no reasonable reading of the handball laws can a player be punished if they have touched the ball despite keeping their arms behind their back, as defenders have recently started doing when blocking shots in the penalty area. But the referees’ interpretive discretion is considerable. Those watching football in different countries will be familiar with this issue. The threshold for being awarded a penalty kick for a foul is considerably higher in the English Premier League, where referees require substantial physical contact, than in the Italian Serie A or Spain’s La Liga, where lighter challenges are also punished. Similarly, there is significant variation in the degree to which players and managers are allowed to protest against a refereeing decision before being cautioned or sent off. This is a direct consequence of the standard-like nature of the applicable football laws.

The divergences are further magnified by the lack of a system of precedent. In law and regulation, standards often turn into rules over time as a consequence of a growing body of jurisprudence or administrative practice (Kaplow 1992, 577).

Once the Supreme Court decides that not looking left and right before crossing the street is negligent as it violates what can be expected from a reasonable person, or an agency finds that products containing a certain chemical compound constitute a danger to human health as they increase the risk of cancer, this decision binds future assessments. Such mechanisms do not exist in football. A referee is, at least *de jure*, not obliged to award a penalty kick for a certain challenge just because another referee has awarded one for the same type of challenge in a match last week. What is more, they are not obliged to do so if they themselves took such a decision in a previous game. Strictly speaking, they are not even bound by their own prior decision in that very match. This means that actions potentially violating standards can be assessed every time afresh (see Schauer 2003). National referee associations such as the Professional Game Match Officials Limited (PGMOL) group, which oversees refereeing in the top competitions in England and Wales, provide an informal counterbalance in this context. They make consistency a condition of being assigned games to officiate, demoting inconsistent referees to lower leagues or striking them from the register altogether. Despite this, referees retain a high degree of flexibility when applying football laws.

3. The Video Assistant Referee I: Reviewable Decisions and Incidents

Football has, in comparison to other sports, been late in embracing technological aids.⁸ Although calls for replay systems have been made since the advent of TV broadcasting, they faced continued opposition at the higher echelons of football governance. It was feared that video technologies were too expensive and unreliable, that they would affect the simplicity and fluidity of football, and, somewhat oddly, that they would rob fans of the joy of debating controversial officiating decisions (see Ryall 2012). Only in the late 2000s and early 2010s the push for reform gained momentum. A series of refereeing errors surrounding the 2010 World Cup – amongst them a goal scored by Thierry Henry in the qualifiers that should have been disallowed due to handball and an incident in the round of 16 where England were denied a goal against Germany although the ball had clearly

passed the line – made it plain that change was necessary. As a result, in 2012 the GLT was adopted, an electronic system that verifies and informs the referee whether a goal has been scored. After gaining experience with the new device and the departure of some members of the “old guard” at FIFA due to a corruption scandal, the ground was ready for a deeper dive into technological aids. In 2018, after extensive testing, the VAR was introduced. It has, since, gradually been put to use in all major domestic and international competitions, including the 2018 men’s and 2019 women’s World Cups, the Champions League, the Copa Libertadores, and the Club World Cup.

When and how the VAR ought to be used is laid down in a special protocol attached to the Laws of the Game. The system is based on the principle of “minimum interference, maximum benefit” (IFAB 2016). At the centre of officiating remains the on-pitch referee (and their assistants, formerly known as “linesmen”), who always must make the initial decision and is responsible for the final call. The VAR, who sits in a special video operation room and has access to live footage of the game from different camera angles, silently checks all decisions made in relation to an exhaustive list of four reviewable actions, but is only supposed to intervene in the event of a “clear and obvious error” or “serious missed incident”. The ultimate decision is to be taken by the referee based on information from the VAR or after undertaking an on-field review on a dedicated screen located on the side of the field.

The decisions and incidents that the VAR can review are limited to: (1) goal/no goal; (2) penalty/no penalty; (3) direct red card; and (4) mistaken identity (for yellow and red cards). Although that might not seem much, this catalogue means that the VAR has a supervisory function for over a dozen of different types of calls. Regarding the issue of goal/no goal, it must review, inter alia, whether the attacking side has committed an offence in the build-up to or scoring of the goal (through a foul, handball, or offside), whether the ball has gone out of play, and whether the ball has crossed the line. In relation to penalties, its tasks include monitoring whether a penalty kick was correctly awarded or not awarded, whether the offence

leading to the penalty was committed inside the penalty area, and whether any other offence has occurred during the build-up. Finally, there is a handful of offences triggering a red card, such as serious foul play and violent conduct, in addition to the issue of punishing the right player.

Some of these reviewable decisions are based on rules and it is here that the VAR has brought the greatest improvements. This notably concerns offside. In the pre-VAR world, incorrect offside calls were a normal occurrence in football games, with studies consistently reporting error rates between 11 and 26%.⁹ While some scholars attributed the problems to optical errors caused by a wrong positioning of the assistant referees (Oudejans et al. 2000), others put forward the so-called “flash lag” hypothesis according to which the human eye has difficulties in identifying the position of two objects if one of them moves (Helsen et al. 2006; Catteeuw et al. 2010). The VAR has dramatically reduced mistakes in this area. Even if Gianni Infantino’s announcement after the 2018 World Cup that offside is “finished” may have been overly optimistic,¹⁰ it is true that, with regard to clear offside, incorrect calls have essentially disappeared. The situation is more complicated for marginal offside, an issue I shall address below.

Similar improvements can be observed in relation to other reviewable rule-based incidents. Cases of mistaken identity were never too frequent in football, but could, at times, have severe consequences, for instance if a player wrongly received a second yellow card and was sent off – they have vanished since the introduction of the VAR. The same goes for errors relating to the location of a foul (inside or outside the penalty area), which determine whether the punishment is a penalty or free kick. Even infractions of recently added rules are being monitored effectively. A 2019 amendment to the Laws of the Game established that goals scored directly from an attacking player’s hand or arm or, immediately after, by their teammate (offensive handball) are prohibited, regardless of whether the touch is accidental or not. It has been put to effect quickly by referees and video assistants across the world.

The VAR system has been much less successful in bringing a sense of “football justice” when it comes to decisions involving standards. Reviews concerning penalties continue to prompt frequent and fierce criticisms from players, managers, and fans, at levels comparable or even higher than pre-VAR.¹¹ Penalties for handball have proven to be one persistent bone of contention. As explained earlier, the handling of the ball by a player from the defending side is not per se prohibited; it must either be “deliberate” or “unnaturally” enlarge their body. Refereeing decisions relating to these requirements regularly end up being disputed. Numerous examples could be cited, but one that has particularly stuck with fans of the German Bundesliga is the home game of Borussia Dortmund against Bayern Munich in the 2019/20 season. Halfway through the second half, a promising shot by Dortmund’s striker Erling Haaland in the penalty area was blocked by Bayern’s centre back Jerome Boateng, who extended his elbow at a 45 degree angle, thus effectively adding a triangle-shaped surface to the right part of his body. To the astonishment of the Dortmund supporters, the referee did not award a penalty – and the VAR did not intervene. What further fuelled their anger was that merely five days later, during Dortmund’s next fixture, their opponent Paderborn was awarded a penalty for what essentially was a carbon copy of the incident, just this time with Borussia’s defender Mats Hummels using his elbow. Again, the VAR did not intervene.

Calls concerning penalty kicks for a foul remain just as disputed. The intensity and type of contact that referees require for awarding a penalty kick substantially differs and the VAR has not managed to remove controversy from this area. Again, plenty of examples could illustrate the point, but one representative incident happened during the Premier League match between Norwich City and Manchester United in October 2019. United’s striker Daniel James was moving into the penalty area at a quick pace, closely followed by a Norwich defender. Both players ended up leaning into each other, trying to get a physical advantage, before going to the ground. It was one of these 50/50-calls or, perhaps more accurately, a 40/60-call, as the video footage appeared to show that the United striker was doing more of the pushing. The referee on the pitch decided that there had been no defensive foul, but the VAR

intervened and, after a lengthy break, advised to award the penalty kick. The intervention sparked much debate, with many arguing that an initially correct decision had been turned into an incorrect one.

Why has the VAR been so effective in improving football officiating in one domain and much less so in the other? The answer has to do with the technology's design and the nature of the football laws it is tasked with monitoring. The VAR, in essence, is an additional referee with an excellent view over what is happening on the pitch. It is, much like the name of the company producing it, a "hawk eye": it captures all the incidents occurring during a match in high resolution, including those which the on-pitch referee has not seen or not seen well. Thus, its main advantage over the head referee, and its core added value to football officiating, is epistemic. This is why the VAR is exceptionally good at policing rule-like offences, which primarily depend on whether certain facts have or have not occurred. Whenever the on-pitch referee does not see a certain action (e.g. because they are looking in a different direction or are at the other end of the field) or has a limited view of it (e.g. due to a player standing in the way), the VAR provides a further, very capable pair of eyes that capture the incident, thus reducing the number of factual errors. (It is for the same reason that the GLT, the other technological aid in football, is widely perceived to have been a success.¹²)

The VAR has no comparable advantage when it comes to evaluative judgments. It is not a more just or fair referee, who is better at interpreting the laws of football – it is only a second opinion, albeit an influential one. Therefore, where a normative decision is at the heart of a call, as is the case in most standard-based offences, the VAR is of limited use. The person sitting in the video operating room, typically a referee themselves, will have a view on whether a certain challenge constitutes "excessive force" or "violent conduct", but this view is not inherently preferable over that of the on-pitch referee. To put it differently, the VAR is no magical device that solves our normative problems and disagreements concerning the interpretation of football laws. And disagreements there are. One of the few studies examining

refereeing performance in relation to fouls (Samuel et al. 2019), the most frequent standard-based call in football, found that there was less than 63% consensus among experienced referees as to whether a certain challenge constituted a foul and only 26% as to whether a yellow or a red card should be given. The VAR cannot change this. It is unable to turn an open-textured football law allowing for different ways of interpretation into a clear and easily applicable norm just by replaying the scene on a screen.

This is not to say that the VAR is of no use for monitoring standard-based incidents. Every standard-based call has a factual component. Where, for instance, the on-pitch referee erroneously believes to have spotted a contact between the defending side's goalkeeper and the attacking team's striker and consequently decides to award a penalty (as e.g. during the 2019/20 Champions League match between Manchester City and Shakhtar Donetsk), the VAR will be able to correct the factual assessment. The same goes for situations in which the referee is unable to see the incident, because they were too far away from the action or a player was blocking the view. Moreover, there can be the occasional serious evaluative error which the VAR will be able to correct. Think, for example, of the above-mentioned situation where a penalty kick is awarded for handball although the defender has kept their hands behind their back. This is the type of call which is so blatantly wrong that nine out of ten referees seeing the incident would have decided otherwise. But these will, as the initial experience with the VAR suggests, be relatively rare.

Viewers of the English Premier League are likely to accept the foregoing depiction of the VAR's problems concerning standard-based calls, but they may object that the system does not fare much better in relation to rule-based calls, especially those concerning offside.¹³ Hardly a week passes in the Premier League without at least one contentious VAR offside decision. The relevant scenes are usually similar. The attacking team's player appears to be at the same level as the opponent's last defender in the moment of the pass, the action continues until the former scores (assistant referees are meant to delay the flag unless the offside is evident), a call is

made (offside or no offside), and then the measuring begins: the VAR's calibrated lines show that the former's knee, chest, or shoulder is a couple of centimetres ahead of the latter's shoulder, or *vice versa*, and the offside decision is upheld or overturned. Call this marginal offside.

Why are marginal offside calls a more prominent issue in England than elsewhere?¹⁴ The reason is two-fold. To begin with, football in the Premier League is faster than in any other domestic or international competition, meaning that the sheer number of offside, both clear and marginal, is higher. Several factors are likely to have contributed to this development: a general increase in the pace of football over the past two decades; the parallel surge of the financial strength of Premier League clubs, which means that they can attract the best talent on the market; and a historical focus in English football on the virtues of speed and attacking, sometimes derided as “kick and rush”.¹⁵ Moreover, and related to this, interruptions to the game cause – due to the cultural expectation towards fast football – greater unhappiness among English spectators than those in other countries, where the sport is played at a slightly more leisurely pace.

The more puzzling and important question is why these ultra-tight offside calls are so contentious. Other than the problem of delay (which may not, in fact, not be as serious as commonly assumed, with preliminary data indicating that the VAR delays matches by a mere 50 seconds on average¹⁶), it appears that the controversy is, partly, fuelled by technological issues and, partly, by broader questions relating to function of football refereeing and the formulation of the offside rule.

Let us look at the technological side first. Collins (2020) has raised concerns about the ability of the VAR to accurately make ultra-tight offside calls.¹⁷ For assessing offside, the position at the moment the ball is played or touched by the attacking team's passing player is relevant. To determine that moment is, in practice, harder than it may seem. The video footage which VAR review is based on is shot at 50 frames per second. Yet, sometimes multiple frames can appear to capture the first

point of contact. The eventual choice of frame will, then, potentially affect the outcome of the offside call. According to one calculation, the error margin connected with choosing one frame over the one directly preceding it is around 19 cm (Lawless 2019). Although this number is in fact likely to be lower (the estimate is based on one of the Premier League's fastest players running at full speed, which is unrepresentative of the vast majority of offside incidents), the point remains a valid one. VAR review has a non-negligible margin of error, which is neither displayed in the video footage nor explained to the public.

Further, Collins has suggested that marginal offside should not be corrected as a matter of principle. Arguing that sports officiating is about “justice not accuracy” (see already Collins 2010; Collins et al. 2016), he recommends rethinking offside, which “should not be seen as a binary decision – one side of a line or the other” but “a matter of human judgement” (Collins 2020, 44). His proposal reads as follows: “only if a player *looks offside* is the player offside”. Others have made similar suggestions, advocating that offside should be reformulated so that it requires there to be “clear daylight” between the attacking player and the opponent’s defender.¹⁸ Both proposals – requiring that a player “looks offside” or has a “clear” advantage – would turn the offside rule into a standard. This, however, is unlikely to reduce the controversy surrounding marginal offside; it is more probable to have the exact opposite effect. If you think that the current levels of dissatisfaction in relation to offside calls are worrying, just imagine what will happen if, in addition to the factual component (is the player ahead of the last defender?), there would be an evaluative one (is the gap sufficiently visible or clear?). The experience in leagues which have experimented with removing calibrated lines from the VAR video footage is telling. The effect was a decrease in both predictability and decision-making accuracy.¹⁹

A simpler and more effective solution might lie in making the technology more transparent by adding error margins of, say, 10 to 15 cm (which could be represented by a thicker calibrated line in the video), an option currently tested in the Dutch Eredivisie. This would kill two birds with one stone. Not only would the

doubts concerning the transparency and reliability of the video analysis be attenuated as possible errors would be laid bare, it would also lead to a situation where only actions in which the attacking side's player is visibly ahead of the last defender are ruled out. In addition, and more ambitiously, a reformulation of the offside law could be considered indeed, but one that maintains its rule-like character. When a rule proves to be repeatedly over-inclusive, it may need tweaking to align it better with its underlying background principle. In the case of offside, this is: no player should have an undue advantage by having a shorter distance to traverse to the goal than the opponent's defenders. Against this backdrop, it might, for instance, make sense to exclude body parts such as armpits and shoulders when assessing offside, and focus exclusively on the players' feet.²⁰ Although the former can be used to score goals (and are, therefore, currently counted when assessing offside), it is only the latter that allow players to cross the pitch.

4. The Video Assistant Referee II: Threshold for Intervention

The VAR protocol stipulates that the video assistant should only step in if there was a "clear and obvious error" or, practically less important, a "serious missed incident". Where the threshold for intervention is set is crucial as it affects both how frequently the VAR will get active and how many of the head referee's decisions will be corrected. In theory, the system could be designed in a variety of ways, but two main options should be distinguished. The first consists in giving the VAR the power to exercise what, in some appellate legal systems, is called *de novo* review. This entails a complete re-assessment of the incident "from scratch", i.e. disregarding any prior refereeing decision made, and would lead to the VAR correcting each and every (reviewable) mistake that occurred on the pitch. The second option is having a system based on deferential review. Here, the VAR would, as a rule, defer to the on-pitch referee's decisions and only step in where a call was deemed evidently wrong. This is the approach that the IFAB has chosen. It reduces not only the number of interventions but also that of reversals, which are limited to blatant errors.

Formally, the VAR is part of the refereeing team and merely “assists” the head referee in their work, akin to the assistant referees standing on the sides of the pitch. In reality, however, it acts more like an external review authority. The VAR is a geographically separate unit, typically located miles away from the stadium and the rest of the refereeing team in a special video operating room. Further, its role is to oversee the mistakes of the other referees. Preliminary evidence indicates that on-pitch referees find it hard to decide against the VAR. This is understandable psychologically given that they are signalled, in front of the entire live and TV audience, that they may have just committed a fundamental and, to make matters worse, “obvious” mistake. From the 50 checks that led VAR to advise the referee to review their decision during the 2017/18 Bundesliga season, 48 led to a change in decision (96%); the corresponding reversal rates for the Portuguese Primeira Liga and the 2018 World Cup are 68% and 82% respectively.

Although, in theory, the “clear and obvious error”-test is meant to apply to all reviewable decisions and incidents, a different practice has gradually emerged. Only in relation to standards is the test adhered to – in relation to rules, all errors are corrected, including minor ones, thus effectively establishing a *de novo* type of review. This two-pronged approach appears to have, *sub silencio*, been embraced by referees and video referees in leagues across the world. In England, it has even been adopted explicitly. In its statement on the application of the VAR, the Premier League (2018) explains that “factual decisions such as offsides, and the issue of whether a player is inside or outside the penalty area, are not subject to the ‘clear and obvious’ test.”

This bifurcated approach, even if contradicting the IFAB’s guidelines, makes intuitive sense. Why should consequential factual errors in football games be tolerated if they can be rectified? Or, to put it more bluntly: if a call is wrong and we can see that it is wrong, why should it stand? It is precisely the VAR’s *forte* to spot and correct factual mistakes. Assuming that the technology is reliable and excessive delays can be avoided, there appears to be little reason to let incorrect

calls stand, especially if they are significant for the outcome of the game.²¹ (A different question is whether avoiding delays should even be given that much weight at the expense of accuracy.²²) The considerable gains in terms of decision-making accuracy and fairness appear to outweigh the limited costs in terms of time and effort (similarly, see Berman 2011).

The more vital element of the two-prong approach, however, is the strict adherence to the “clear and obvious error” test in relation to standards, especially when it comes to the evaluative component of a call. Here, the VAR has no inherent advantage over the on-pitch referee. Mistakes still can happen: factual errors are possible as are occasional blunders in how a football law is applied. But, outside these situations, a healthy dose of restraint on the VAR’s part is warranted. Otherwise one risks substituting one reasonable interpretation of a football law (that of the on-pitch referee) for another (that of the VAR).²³ The first VAR season in the Bundesliga provides a cautionary tale in this respect. After a turbulent first couple of months with the new technology, an internal communication of the German football association was leaked to the public, showing that referees had been advised to look for VAR guidance not only in cases of clear errors but also when they simply had doubts about how to decide. The result was a high number of “advisory opinions”, which led to initial reasonable calls being overturned. The disclosure prompted a great deal of backlash and, as a result, the guidelines were changed back to the “clear and obvious error” test. This, in turn, reduced the number of VAR interventions, while increasing the overall satisfaction with the system (Kolbinger 2019, 233-236).

Hence, there are plausible reasons for the evolving practice of subjecting rule-based decisions to a more lenient and standard-based decisions to a more stringent test. However, as always, context matters. In leagues marked by a greater-than-average distrust in football referees, for instance due to recurring corruption scandals, adopting a less demanding test for VAR interventions can make sense. It will introduce a further layer of protection that will render match-fixing more difficult.

Where, to the contrary, we have high numbers of minimal rules infractions the policing of which takes a great toll on the fluidity of the game, it might be justified to impose a more stringent threshold.²⁴ The problems with marginal offside in the Premier League are an example for an environment where raising the bar for VAR interventions – or, in the words of the IFAB’s general secretary, not being “too forensic” (BBC Sports 2019) – could be an appropriate solution (if the reforms discussed above are not implemented). The League’s decision to adopt *de novo* review for offside calls is certainly counterproductive in this regard.

In addition to affecting *when* the VAR should intervene, the rules/standards distinction is also helpful for considering *how* this intervention should take place. The VAR protocol distinguishes between two types of video assistance. For “factual decisions”, it states that a VAR-only review will usually be appropriate, meaning that the on-pitch referee does not have to watch the incident themselves but can rely on the information provided by the video assistant. Such decisions include the position of an offence or player (for offside and penalty), the point of contact (for handball and foul), and questions such as whether the ball has gone out of play. For “subjective decisions”, like the intensity of a foul and handball considerations, an on-field review is recommended. This involves the referee themselves taking another look at the incident on the side-pitch monitor.

Some leagues – among them, at least initially, the Premier League²⁵ – have chosen to reduce the number of on-field reviews and made all or most reviews VAR-only. If what has been argued throughout this article is correct, the IFAB is right in insisting that these associations bring their national guidelines in line with its VAR protocol. VAR-only reviews are a good idea when it comes to factual decisions, most of which revolve around the application of rules or the factual component of a standard-based call. Letting the VAR determine the issue will usually be faster and yield as accurate, in some cases even more accurate results – research suggests that physical factors like exhaustion can affect the ability to decide correctly (Catteuw et al. 2010) – than making the referee review the incident on a pitch-side

monitor. The situation is different for subjective decisions, which all concern standard-based calls. The VAR protocol recommends an on-field review here precisely to avoid the aforementioned risk of substituting the VAR's assessment for that of the referee. Allowing the latter to re-watch the incident themselves ensures that ultimately the central and often difficult normative decisions will, as intended, be made by the on-pitch referee, not the video assistant.

5. Conclusion

Just like legal norms, football norms can be divided into rules and standards. The choice between the two does not only affect the way in which football laws are enforced in general, it has repercussions for the use of technological aids. The VAR has shown to be a valuable tool for enforcing rule infractions, while contributing comparatively little to strengthening "football justice" when it comes to decisions turning on standards. This varied impact is not because the technology is insufficiently developed, it is a consequence of the nature of the football laws it is tasked with adjudicating on. The proverbial "hawk eye", the VAR is able to reduce epistemic problems in football officiating. This is why it performs so well when it comes to rule-like offences, which require factual assessments only. It has no comparable advantage in relation to evaluative judgments, which explains the persisting difficulties the system faces when dealing with standard-like offences such as fouls and handball violations.

The findings of this article have relevance beyond the "beautiful game". The rules/standards distinction can be found in most sports, many of which have introduced, or considered introducing, technological aids. The present analysis highlights the importance of looking not just at the supply side but the demand side when doing so. As important as the design and setup of the aids themselves is the nature of the tasks they are meant to fulfil. When setting out the competences of devices like the VAR, it is crucial to consider the environment they are applied in and the norms are expected to enforce. This reflective process might reveal that the

technology's potential is high in some areas and surprisingly low in others, insights that will inform decisions as to whether a given system should be established to begin with and how it can, once in place, be sensibly reformed.

Notes

1. But see Nlandu 2011, 30-31.
2. This is even true of the aforementioned speed limits: in many countries, driving faster than allowed is illegal but minimal infractions, e.g. driving less than 3 miles per hour over the limit, will not be penalised.
3. The fact that the distinction is not clear-cut does not mean that there are not significant differences between the two types concerning their clarity and predictability of application as well as the scope of interpretive discretion which officials have (see remainder of section).
4. Similarly, for tennis, see Berman 2011b, 1361 et seq.
5. "Passive" offside is formulated in a more standard-like manner: it requires the player who is in an offside position to interfere with an opponent, for instance, by preventing them from playing or being able to play the ball by "clearly obstructing the opponent's line of vision" or "making an obvious action which clearly impacts on the ability of an opponent to play the ball" (Law 11.2).
6. It includes attempts to deceive the referee (e.g. by diving), committing an offence that interferes with a promising attack of the opponent, and "showing a lack of respect for the game"; see IFAB 2020, Law 12(3).
7. The "six second rule" whereby the goalkeeper must not take longer than six seconds to release the ball, which is meant to prevent time play, has, due to underenforcement, effectively become a standard prohibiting exceedingly long tardiness.
8. Instant replays were introduced in 1986 by the American National Football League (but temporarily suspended between 1992 and 1999) and in 2002 by the National Basketball Association. The Hawk-Eye system has been used since 2006 in tennis and 2001 in cricket, which had already experimented with a variety of other technological aids already since the early 1990s.
9. Helsen et al. 2006; Oudejans et al. 2005; Catteuw et al. 2010; Hüttermann et al. 2017.
10. See www.reuters.com/article/us-soccer-worldcup-fifa-infantino-idUSKBN1K31IG.
11. The Twitter account "VAR-watch", which monitors the technology's performance in the German Bundesliga, has rated 13 VAR decisions during the 2019/20 season as "questionable", 8 of which concern penalties.
12. This is not to say that there have not been problems. The French Football League suspended the use of a GLT system developed by GoalControl in 2018 after a series of errors, before signing a contract with market leader Hawk-Eye. But even Hawk-Eye's technology occasionally malfunctions as the Premier League match between Sheffield United and Aston Villa in June 2020 showed. GLT failed to detect a goal for Sheffield although Villa's goalkeeper had, after catching a ball from a cross, carried it over the goal line. Hawk-Eye explained that this was the result of the level of "occlusion" by the players involved, which had not been reached in the 9,000 previous matches during which the system was applied (Ostlere 2020). Football commentators were quick to point out that the VAR could – and should – have intervened to correct the mistake. These incidents should not detract from the fact that the GLT has worked remarkably well in the vast number of games in which it has been applied, leading to correct decisions in critical situations, some of which have decided titles races, such as the goal-line clearance by John Stones in the match between Manchester City and Liverpool in the 2018/19 season, where GLT correctly signalled that the ball was not fully over the goal line, with only 11 millimetres missing.
13. See Collins 2020, 9.
14. Addressing the growing frustration in the Premier League, Gianni Infantino has noted that "In Italy, in Spain, in Portugal, in Germany this offside discussion is not existing on the VAR." See

uk.reuters.com/article/uk-soccer-ifab/fifa-chief-infantino-says-open-to-changes-in-offside-law-idUKKCN20M2X7.

15. See Wilson 2013. This does not mean that the phenomenon will remain a predominantly English one. The modern game becomes faster across the world and is, for tactical reasons, played on an increasingly confined space on the pitch, which is likely to contribute to the problems described below.

16. Figures for the 2019/20 Premier League season: see www.premierleague.com/news/1293321. These confirm earlier findings from the Bundesliga (Deutsche Fussball Liga 2018), where VAR caused a delay of 60 seconds per game on average during its first season.

17. Collins 2020.

18. See uk.reuters.com/article/uk-soccer-ifab/fifa-chief-infantino-says-open-to-changes-in-offside-law-idUKKCN20M2X7.

19. During the first half of the 2017/18 Bundesliga season, the German Football Federation decided to not employ calibrated lines for VAR offside reviews and, instead, relied on simple video footage of the scene, leading to a situation where only if a player “looked offside” he was offside. The result was several incorrect and many contentious offside calls (Kolbinger 2019), triggering much dissatisfaction among players and commentators. In reaction to the criticisms, the calibrated lines were introduced during the second half of the season, eliminating much of the controversy.

20. As of the 2020/21 season, the definition of handball has changed: the boundary between the arm and the shoulder (which is the boundary between handball and no handball) is now defined as the bottom of the armpit. This, on the one hand, means that players can score goals with the side of their shoulder. On the other hand, it also has the effect that players are more easily in an offside position, as their shoulder counts as a relevant body part for the purpose of the offside decision. For an illustration of the problem see www.bbc.co.uk/sport/football/54855596 (discussing the ruling out of a Patrick Bamford goal “for pointing” in a match between Leeds United and Crystal Palace).

21. A related question, which is bracketed in this paper, is whether this means that the VAR’s competence catalogue should be expanded beyond the four decisions/incidents that are currently reviewable. In most sports, technological aids are used to correct consequential refereeing errors only, not all errors. This limitation is the result of a cost/benefit analysis: while there is a lot to be gained from correcting a mistake that has a significant impact on the outcome of the game (e.g. goal/no goal), there are few, if any, benefits attached to correcting minor infractions with no genuine impact (e.g. simple foul in the middle of the pitch). As argued by Berman, video review systems should be designed so that they “maximize error correction up to the point at which the marginal cost exceeds the marginal benefits” (2011, 1693; cf. Nlandu 2011). This does not mean that the current list of reviewable decisions/incidents strikes the ideal balance between the two, but it makes a *prima facie* case for not reviewing every refereeing mistake.

22. Berman 2011, 1703-1706.

23. This is a major difference between football and other sports where technological aids are, primarily or even exclusively, used to enforce rule-like norms, such as American football; see Berman 2011.

24. Having different VAR regimes in different countries or leagues is in tension with the principle of universality underlying the Laws of the Game. But the IFAB has, over the past years, warmed up to the idea of differentiated and context-specific arrangements, e.g. in lower-tier and youth competitions.

25. As of the 2020/21 season, the Premier League is following the IFAB’s VAR Protocol on the use of on-pitch monitors.

Acknowledgements

I wish to thank Jo Murkens, Guilherme Sampaio, Joe Tomlinson, Bosko Tripkovic, Steve Weatherill, and Jack Williams as well as the two anonymous reviewers for their stimulating comments.

References

- BERMAN, M. 2011. Replay. *California Law Review* 99 (6): 1683-1744.
- BBC SPORT. 2019. VAR “Should Not Be Too Forensic” – Football Law-Makers Set to Issue Guidance. Available at www.bbc.co.uk/sport/football/50944416 (accessed 16 November 2020).
- CATTEEUW, P., B. GILIS, J. WAGEMANS and W. HELSEN. 2010. Offside Decision Making of Assistant Referees in the English Premier League: Impact of Physical and Perceptual-Cognitive Factors on Match Performance. *Journal of Sports Sciences* 28 (5): 471-481.
- COLLINS, H. 2010. The Philosophy of Umpiring and the Introduction of Decision-Aid Technology. *Journal of the Philosophy of Sport* 37 (2): 135-146.
- COLLINS, H., R. EVANS, and C. HIGGINS. 2016. *Bad Call: Technology’s Attack on Referees and Umpires and How to Fix It*. Cambridge: MIT Press.
- COLLINS, H. 2020. This is not a penalty! What’s gone wrong with technology and football in the age of VAR? Draft Working Paper, available at sites.cardiff.ac.uk/harrycollins/files/2020/01/Video_Assistant_Referee_VAR__judgement_versus_measurement.pdf (accessed 16 November 2020).
- DEUTSCHE FUSSBALL LIGA. 2018. Video-Assistent: Hinrunden-Bilanz 2018/19. Available at media.dfl.de/sites/2/2019/01/Video-Assistent-Hinrunden-Bilanz-2018-19.pdf (accessed 16 November 2020).
- DWORKIN, R. 1977. *Taking Rights Seriously*. Cambridge: Harvard University Press.
- EHRlich, I. and R. POSNER. 1974. An Economic Analysis of Legal Rulemaking. *The Journal of Legal Studies* 3 (1): 257-286.

HELSEN, W., B. GILIS and M. WESTON. 2006. Errors in Judging “Offside” in Association Football: Test of the Optical Error Versus the Perceptual Flash-Lag Hypothesis. *Journal of Sports Sciences* 24 (5): 521-528.

HÜTTERMANN, S., B. NOËL and D. MEMMERT. 2017. Evaluating Erroneous Offside Calls in Soccer. *PLoS ONE* 12(3): e0174358.

INTERNATIONAL FOOTBALL ASSOCIATION BOARD (IFAB). 2016. Video assistant referees (VARs) experiment: Protocol (Summary).

INTERNATIONAL FOOTBALL ASSOCIATION BOARD (IFAB). 2020. Laws of the Game 2020/21. Available at www.theifab.com/log-documents (accessed 16 November 2020).

KAPLOW, L. 1992. Rules Versus Standards: An Economic Analysis. *Duke Law Journal* 42 (3): 557-629.

KOLBINGER, O. 2019. VAR Experiments in the Bundesliga. In *The Use of Video Technologies in Refereeing Football and Other Sports*, edited by Armanteros, Benítez and Betancor. New York and London: Routledge: 228-244.

KOROBKIN, R. 2000. Behavioral Analysis and Legal Form – Rules vs Standards. *Oregon Law Review* 79 (1): 23-60.

LAWLESS, J. 2019. Calculations Show How VAR Is Not Suited for Tight Offside Calls. *Sport Bible*. Available at www.sportbible.com/football/news-calculations-show-how-var-is-not-suited-for-tight-offside-calls-20190922 (accessed 16 November 2020).

NLANDU, T. 2011. One Play Cannot be Known to Win or Lose a Game: a Fallibilist Account of Game. *Sport, Ethics and Philosophy* 5(1), 21-33.

NLANDU, T. 2012. The Fallacies of the Assumptions Behind the Arguments for Goal-Line Technology. *Soccer, Sport, Ethics and Philosophy* 6 (4): 451-466.

OSTLERE, L. 2020. Hawk-Eye Blames Freak “Occlusion” for Goal-Line Technology Error During Aston Villa vs Sheffield United. *The Independent*. Available at www.independent.co.uk/sport/football/premier-league/aston-villa-sheffield-united-goal-line-technology-hawk-eye-norwood-nyland-var-a9571901.html (accessed 16 November 2020).

OUDEJANS, R., F. BAKKER, R. VERHEIJEN, J. GERRITS, M. STEINBRÜCKNER, and P. BEEK. 2005. How Position and Motion of Expert Assistant Referees in Soccer

- Relate to the Quality of Their Offside Judgements During Actual Match Play. *International Journal of Sport Psychology* 36: 3-21.
- PREMIER LEAGUE. 2018. VAR – Frequently Asked Questions. Available at www.premierleague.com/news/1293321 (16 November 2020).
- RUSSELL, J.S. 1999. Are Rules All an Umpire Has to Work With? *Journal of the Philosophy of Sport* 26 (1): 27-49.
- RYALL, E. 2012. Are There any Good Arguments Against Goal-Line Technology? *Sport, Ethics and Philosophy* 6 (4): 439-450.
- SAMUEL, R.D., Y. GALILY, O.G.E. SHARONI, and G. TENENBAUM. 2019. A Decision-Making Simulator for Soccer Referees. *International Journal of Sports Science* 14(4): 480-489.
- SCHAUER, F. 1991. *Playing by the Rules: A Philosophical Examination of Rule-Based Decision-Making in Law and in Life*. Oxford: Clarendon Press.
- SCHAUER, F. 2003. The Convergence of Rules and Standards. *New Zealand Law Review* 3: 303-328.
- SCHLAG, P. 1985. Rules and Standards. *UCLA Law Review* 33 (2): 379-430.
- SIMÒN, J.A. 2019. Between Scepticism and the Long-Awaited Quest for “Football Justice”. In *The Use of Video Technologies in Refereeing Football and Other Sports*, edited by Armanteros, Benítez and Betancor. New York and London: Routledge: 2-22.
- SULLIVAN, K. M. 1992. Foreword: The Justices of Rules and Standards. *Harvard Law Review* 106 (1): 22-123.
- WILSON, J. 2013. *Inverting the Pyramid: The History of Soccer Tactics*. London: Orion.