Routledge
Taylor & Francis Group

Journal of Sports Sciences

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/rjsp20

Video assistant referees (VAR): The impact of technology on decision making in association football referees

Jochim Spitz, Johan Wagemans, Daniel Memmert, A. Mark Williams & Werner F. Helsen

To cite this article: Jochim Spitz, Johan Wagemans, Daniel Memmert, A. Mark Williams & Werner F. Helsen (2021) Video assistant referees (VAR): The impact of technology on decision making in association football referees, Journal of Sports Sciences, 39:2, 147-153, DOI: 10.1080/02640414.2020.1809163

To link to this article: https://doi.org/10.1080/02640414.2020.1809163

JOURNAL OF SPORTS SCIENCES
2021, VOL. 39, NO. 2, 147–153
https://doi.org/10.1080/02640414.2020.1809163

Routledge
Taylor & Francis Group

SOCIAL AND BEHAVIORAL SCIENCES

Video assistant referees (VAR): The impact of technology on decision making in association football referees

Jochim Spitz[a], Johan Wagemans[b], Daniel Memmert[c], A. Mark Williams[d] and Werner F. Helsen[a]

[a]Department of Movement Sciences, Laboratory of Perception and Performance, Movement Control and Neuroplasticity Research Group, University of Leuven (KU Leuven), Leuven, Belgium; [b]Department of Brain & Cognition, Laboratory of Experimental Psychology, University of Leuven (KU Leuven), Leuven, Belgium; [c]Institute of Exercise Training and Sport Informatics, German Sport University Cologne, Köln, Germany; [d]Department of Health and Kinesiology, College of Health, University of Utah, Salt Lake City, UT, USA

ABSTRACT

The use of technology has been proposed to improve decision-making in sport officials. The implementation of the video Assistant Referee (VAR) in association football is one example of how technology can be used to assist decision making, although its impact remains unknown. In 2195 competitive football matches across 13 countries, the VAR conducted 9732 checks for potential match-changing incidents, with the median duration of a check being 22 seconds. The checks resulted in a total of 795 reviews, with a median duration of 62.0 s for on-field reviews (N = 534) and 15.0 s for VAR-only reviews (N = 261).We report that the predictive odds for making the correct decision after VAR intervention were significantly higher than for the initial referee's decision, with accuracy increasing from 92.1% to 98.3%. Findings have implications for the current debate about the introduction of technology in association football and may help set guidelines regarding the use of technology across other sports and professional domains.

ARTICLE HISTORY
Accepted 23 July 2020

KEYWORDS
Video technology; sport officials; judgement; soccer

Over the last decades, athletes and coaches have become obsessed with crossing boundaries, setting new records, and getting better, faster and stronger (Coutts, 2016). The development of evidence-based practices and innovations has become a necessity to be successful in elite professional sports. Sport officials have the important responsibility of managing sport competitions at the highest level of play in a fair and accurate way (D. R. Mascarenhas, Collins, Mortimer et al., 2005). Their performance might influence the success of athletes and teams with huge social and economic impact. Over the past few years, governing bodies have been ensuring that referees, like athletes, are provided with the best support. In particular, officials have been allowed to use advanced technological tools in almost every competitive sport. For instance, ball-tracking cameras are used in tennis, association football and cricket to analyse the ball's trajectory and verify whether the ball was in-play, crossed the goal line or a batter should be given out. Moreover, officials in sports like basketball and American football have access to instant replays of situations to review their decisions. These technological innovations help support decision making in officials, increasing accuracy (Leveaux, 2010). In association football, the laws of the game 2018–2019 included the introduction of the video assistant referee (VAR). We investigate the introduction of the VAR in association football and the impact it has had on decision accuracy and the temporal flow of the game.

In association football, referees are required to make rapid decisions while considering several sources of information (D. R. Mascarenhas, Collins, Mortimer et al., 2005; J. Spitz et al., 2016). Scientists have highlighted the possible biasing influences that can occur across different steps in the decision-making process (Plessner & Haar, 2006). An incorrect decision can arise because the referee simply did not see the incident (misperception). Moreover, it appears that sport officials include the context of the situation to make a judgement. The laws of the game leave room for interpretation and researchers have shown that referees may be influenced by crowd noise (Nevill et al., 2002), home advantage (Unkelbach & Memmert, 2010), and previous decisions (Plessner & Betsch, 2001). Other contextual dimensions such as a team's aggressive reputation (Jones et al., 2002), or physical appearance (Van Quaquebeke & Giessner, 2010) can play a role and bias the decisions of referees. The decisions made by referees are therefore not 100% correct and the VAR has been introduced to correct clear and obvious errors for possible match-changing incidents, such as goals, penalty decisions, direct red cards and mistaken identity. The VAR can intervene when an erroneous decision results from one or more steps in the decision-making process (Plessner & Haar, 2006): attention and perception (e.g., when a serious incident in the penalty area was missed or the referee did not see whether the ball crossed the goal line or not); information processing and categorization (e.g., when a serious foul play was wrongly categorized as a yellow card instead of a red card); and the behavioural response (e.g., when an incident was well perceived and categorized but the yellow card was given to the wrong player, which is labelled as mistaken identity).

According to the protocol, the VAR, assisted by a replay operator, checks all video feeds and replays of match-changing situations. If a check reveals that the referee or

CONTACT Jochim Spitz jochim.spitz@kuleuven.be Department of Movement Sciences, Laboratory of Perception and Performance, Movement Control and Neuroplasticity Research Group, University of Leuven (KU Leuven), Leuven B – 3001, Belgium

assistant referee made a clear and obvious error, the VAR recommends that the referee reviews/changes the initial decision. A factual decision, for example, an offside incident before a goal, can be changed based only on the information from the VAR. This is called a VAR-only review. For subjective decisions such as red cards, the main referee can review the footage directly on a monitor near the pitch before making a final decision (on-field review).

The use of technology can improve the accuracy of decisions for situations that require temporal (exact moment) and spatial (where is the ball, where is the player?) precision. For example, previously researchers have reported the usage and success rate of ball-tracking technology in tennis and other sports (Kolbinger & Lames, 2017). Carboch et al. (2016) showed that 27% of the challenges made in top tennis tournaments were line umpire errors with an average error size of 33.2 mm. These errors can be corrected by technology. Moreover, Myint et al. (2015) showed that the possibility of misjudgement in a table tennis game can be reduced by means of technology which captures the position of the ball and table. This approach can be generalized to other clear-cut situations in sports where errors can occur as a result of limitations in human perception; whether a ball hits the foot in field hockey or whether a player is in an offside position at the moment of the pass in association football.

However, it remains unclear to what extent video replays can help referees during a match when decisions require rule interpretation and application (Helsen et al., 2019). The VAR in association football can intervene in four situations which are not clear-cut. In this study, our first aim was to investigate the impact of the VAR on referees' performances (i.e., whether or not the interventions of the VAR improve the accuracy of referees' decisions). Previously, researchers have shown that referees can benefit from slow-motion video, repeated viewings and different viewing perspectives (Pizzera et al., 2016; J. Spitz et al., 2017). On the other hand, researchers have shown that slow motion can make a particular incident or offence seem more pre-meditated than it actually might have been and can, for example, change the disciplinary sanction from a yellow card into a red card (Caruso et al., 2016; J. Spitz et al., 2018). Such findings help provide guidelines regarding the use of slow motion within the VAR protocol. In the current study, we analysed the initial and final decisions (after VAR intervention) of the referees on the field of play and predicted that referees would make more accurate decisions after the intervention of the VAR compared to before.

A secondary aim was to examine the number and duration of interventions by the VAR per match. Technological advancements need to fit the demands of the sport since the typical "flow of the game" can be negatively impacted (Ryall, 2012; Svantesson, 2014). The laws of the game restrict the interventions of the VAR to match-changing incidents: goals; penalty decisions; direct red cards; and mistaken identity. Moreover, VARs only need to intervene in cases of clear and obvious errors and previous published reports suggest that the accuracy scores of referees for a selection of incidents during real games range from 64% to 77% (Gilis et al., 2006; D. R. Mascarenhas et al., 2009). Given the low number of potential reviewable decisions per match (Boyko et al., 2007), and an expected high accuracy score for the referee's initial decisions, we predicted that the number of interventions for clear and obvious errors would be limited with a maximum benefit in terms of accuracy. This prediction would be reflected in a low number of reviews across all matches and an increase in decision accuracy after VAR intervention.

Methods

Data collection

After approval by the ethics committee of the Katholieke Universiteit Leuven in Belgium (G-2015 04 218), data were collected from 2195 matches across 13 national football associations (Australia, Belgium, China, Czech Republic, England, France, Germany, Italy, The Netherlands, Poland, Portugal, South Korea and USA). All countries organized competitive matches employing the VAR during the 2016/2017 and/or 2017/2018 season (from 1st of January 2017 until 1st of June 2018). The use of the VAR was only permitted in competitive matches where the national football association fulfilled all the VAR protocol and implementation requirements as set out in the VAR handbook of the International Football Association Board (the IFAB). The VAR operational protocols are the same for each national football association (Fédération Internationale de Football Association, 2019).

Before introduction of the VAR in the respective competition, a conference call was set up and the VAR project managers, appointed by the competition organizers, received a detailed document with all the relevant information related to data collection. The **13** VAR project managers received an individual login to an online platform for data collection. This online platform was specifically developed for uploading videos with accompanying questionnaires in a secure environment (www.thirdlight.com; www.surveygizmo.com). The project managers were required to upload the video clips from all the checks and reviews within 72 hours after each match that was played with VAR. A detailed description of the definition of a check and review was provided to the project managers to be sure that all the relevant clips were uploaded. After uploading the video clips, the project managers provided general match information and all the relevant information related to the checks/reviews via a standardized online questionnaire. This questionnaire consisted of 24 questions related to the accuracy, duration and impact of the checks and reviews. To ensure that all data were provided in a correct and consistent way by the VAR project managers of all countries, the information from the questionnaires was verified and checked weekly against the information visible in the uploaded video clip and general match data available online (www.wyscout.com).

The members of the national referees committee were asked to independently determine the reference decision for each situation according to the laws of the game (Fédération Internationale de Football Association, 2017, 2018). The referees committees varied in size, but each committee had a minimum of 5 and a maximum of 10 members. Decision accuracy was only calculated for those decisions with complete agreement by the experts on the referees committee. The laws of the game leave room for interpretation and therefore certain

situations for which 100% agreement was not reached by the referees committee members, were classified as "grey zone" situations. These grey zone situations were checked by the Technical Director of the IFAB to ensure consistency across national associations and they were not included in the accuracy analysis.

At least 80% of the national referees' committee members must have at least five years of experience as refereeing specialists and their decisions can be considered as the golden standard for refereeing decisions within a country. When introducing the VAR in the respective competition, the national referees committees were reminded of the guideline to only use slow-motion video for the determination of ball out of play (including goal/no goal), "point of contact" for physical offences and handball. Normal speed should be used to evaluate the "intensity" of an offence or to decide if a handball was "deliberate" (Fédération Internationale de Football Association, 2019).

Dependent variables and statistical analysis

All checks and reviews related to the four categories described in the VAR handbook were analysed: goals; penalty decisions; direct red cards; and mistaken identity. Goal situations were only included in the analysis if the referee had to make a decision in the build-up to the goal for a potential offside position, a ball in/out of play, or an offence by the attacking team.

Decision accuracy

The correctness of the decision was determined separately for the initial decision of the referee (initial decision accuracy) and the final decision after possible intervention of the VAR (final decision accuracy). Accuracy scores were calculated as the total number of correct decisions (in percentage), that is, decisions that were in correspondence with the reference decision of the national referees committee.

Duration of checks and reviews

The check duration relates to the time needed by the VAR team to check the incident and determine whether a review needs to take place. According to the VAR protocol, the review process commences once the main referee clearly indicates this by outlining the shape of a TV screen with both hands and ends when the final decision is given by the referee. The referee has the option to do a VAR-only review, where the final decision is based only on the communication with the VAR, or an on-field review, where the main referee reviews the footage on a monitor near the pitch. The guideline in the VAR protocol is that on-field reviews should not be needed for factual decisions such as the position of an offence or player (e.g., offside), point of contact on the body for handball and ball out of play in the build-up to a goal. The review duration was determined separately from the check duration.

Statistical analysis

We examined the odds for accurate initial and final decisions using a logistic regression model. In this analysis, each situation contributed both an initial and final (after potential intervention of the VAR) decision of the referee. The decisions were not statistically independent because referees had more than one decision. To account for this latter factor, we used a random effects model, with a random term for the referee, fitted using the NLMIXED procedure with the use of the Dual Quasi-Newton optimization algorithm and Gaussian quadrature in Statistical Analyses Software 9.4 (SAS Institute Inc., Cary, NC):

$$log\left(\frac{\mu}{1-\mu}\right) = \alpha + \beta * decision + b$$

In this mixed model, we have one fixed effect (β) for the type of decision (initial/final) and one random effect (b) for the specific referee. Based on the model, the proportions of the correct initial and final decisions were estimated. Finally, descriptive statistics (frequency, median and interquartile range (IQR)) were used for the number and duration of the checks and reviews.

Results

Decision accuracy

There were 9732 checks in total across 2195 matches. Altogether, 638 situations were classified as grey zone areas, that is, decisions for which there is no clear reference decision (i.e., more than one decision could be supported). There were 99 reviews for grey zone incidents and the referee changed the initial decision 44 times (7% of all grey zone incidents).

The referee's initial decision was correct in 8376 of the 9094 clear situations, yielding a decision accuracy of 92.1%. After intervention of the VAR, 8942 of 9094 situations were correct, yielding an accuracy score of 98.3%. Table 1 lists the random effect model parameters, their maximum likelihood estimates, standard errors and inferential statistics.

The parameter estimates indicate that there was a significant effect for the type of decision ($t(250) = -17.98$, $p < 0.0001$), such that the predictive odds to make the correct decision "final" (i.e., after a possible intervention of the VAR), was significantly higher than the initial decision. The odds ratio for success for the initial decision compared to the final decision was 0.19 (95% CI [0.16–0.23]).

In 585 out of the 718 incidents with incorrect initial decisions (81.5%), the decision was reviewed and 577 of these decisions were corrected. In 111 out of the 8376 incidents with correct

Table 1. Random effect model parameter estimates.

Parameter	Estimate	SE (95% CI)	t -value	p	Gradient	OR
α	4.1375	0.0927 (3.9549; 4.3200)	44.64	<.0001	−0.00050	
β	−1.6360	0.0910 (−1.8152; −1.4568)	−17.98	<.0001	−0.00068	0.1948
b	−1.2502	0.2211 (−1.6855; −0.8148)	−5.66	<.0001	0.000019	

initial decisions (1.3%), the decision was reviewed and 11 of these were overturned into an incorrect decision.

Number and duration of the checks and reviews

The median duration of the 9732 checks (4.4 checks per match on average) was 22.0 seconds (IQR = 15.0 s – 35.0 s). The median duration of all checks during a match was 110.0 seconds (IQR = 60.0 s – 171.0 s). Altogether, 795 of these checks resulted in a review (0.36 reviews per match on average), more specifically 534 were on-field reviews with a median duration of 62.0 s (IQR = 48.0 s – 86.8 s) and 261 were VAR-only reviews with a median duration of 15.0 s (IQR: 3.0 s – 32.0 s). There were 1544 matches without a review (70.3% of all matches); 530 matches with only 1 review (24.2% of all matches); 103 matches with 2 reviews (4.7% of all matches); 15 matches with 3 reviews (0.7% of all matches); 2 matches with 4 reviews (0.1% of all matches) and 1 match with 6 reviews (0.1% of all matches). A distribution of the checks and reviews over the different categories of incidents (in %) is provided in Table 2. Red card incidents had the highest proportion of checks (39.3%), followed by penalty incidents (33.4%), goals (27.1%) and mistaken identity (<1%). Penalty incidents (43.9%) had the highest proportion of reviews, followed by goals (32.5%), red card incidents (22.5%) and mistaken identity (1.1%)

In case of a review, the referee has the opportunity to change the initial decision. There were 76 extra penalties (164 penalties awarded; 88 penalties cancelled), 126 extra red cards (132 red cards awarded; 6 red cards cancelled) and 114 fewer goals (61 goals awarded; 175 goals cancelled) due to the interventions of the VAR. An overview of the decision changes based on the interventions of the VAR is presented in Table 3.

Table 2. Number of checks/reviews per category of the incident.

Category	Checks (%)	Reviews (%)
Direct red card	3820 (39.3)	179 (22.5)
Penalty decision	3251 (33.4)	349 (43.9)
Goal	2637 (27.1)	258 (32.5)
Mistaken identity	24 (0.2)	9 (1.1)

Table 3. Number of decision changes after VAR interventions for the different categories of incidents.

	Awarded (#)	Cancelled (#)
Penalty	164	88
Direct red card	132	6
Goal	61	175

Discussion

The VAR was introduced in association football in 2018 to help correct clear and obvious decision errors for possible match-changing incidents, namely, goals, penalty decisions, direct red cards and mistaken identity. In this study, we examined whether the interventions employing VAR improved the accuracy of decisions. Previously, researchers have shown that decision accuracy for clear-cut situations can be improved in sports where errors can occur as a result of limitations in human perception (Carboch et al., 2016; Kolbinger & Lames, 2017; Myint et al., 2015). Our results show, for the first time, using empirical data that there is a significant increase in decision accuracy after VAR intervention for situations that require rule interpretation. The decision accuracy improved from 92.1% to 98.3%. When looking at the in-game decision accuracy scores of referees for specific incidents, there is a range from 64% to 77% (Gilis et al., 2006; D. R. Mascarenhas et al., 2009). The accuracy scores reported are higher compared to previous studies and this is probably due to the fact that all potential match-changing incidents were considered, whereas previously researchers used a specific selection of more difficult incidents. Moreover, referees and VARs in this study were active at the highest levels of their national competition, so could be classified as super-elite officials. The reference decisions in this study were determined by the national referees committee, which can be seen as the golden standard for referees in a country. Given the international scope of the study, the authors were not able to control whether these reference decisions were determined independently by the members of the national referees committee. This may be considered as a limitation of the study. We want to acknowledge the importance of a correct judgement of accuracy of decisions for future research. When researchers do not have access to the referees committees' reference decisions, we advise that they use the consensus decisions of another carefully selected expert panel for the determination of accuracy.

Since there remains a human element in the judgement process, it appears to be impossible to eliminate all clear and obvious decision-making errors and achieve 100% accuracy, even with the use of the VAR (Collins, 2010; Nlandu, 2012; Royce, 2012). Several possible biases can occur in the decision-making process of referees (Jones et al., 2002; Nevill et al., 2002; Plessner & Betsch, 2001; Unkelbach & Memmert, 2010; Van Quaquebeke & Giessner, 2010). Furthermore, a total of 638 situations that were classified as grey zone situations were not included in our analyses. For these situations, there was no clear reference decision and thus no clear correct or incorrect decision. In these cases, decisions can be given within the "spirit or context of the game" and previously published reports have shown that referees apply a "game management strategy" (Balmer et al., 2007; D. R. Mascarenhas et al., 2002; Unkelbach & Memmert, 2008). According to the protocol, VARs should only intervene for clear and obvious errors. In line with the "take the first" heuristic (Johnson & Raab, 2003), referees are therefore reminded to stick to their initial decision for grey zone situations. The laws of the game are evolving, but assessment of particular incidents will remain "subjective" and will continue to cause discussion. With the introduction of the VAR, there might be a tendency to formulate more objective decision criteria, such as is the case for the handball rule (Fédération Internationale de Football Association, 2019).

Overall, refereeing is still considered a "practice-poor" domain given the low number of matches and the limited amount of crucial decisions referees have to make (Catteeuw et al., 2009; Helsen et al., 2019). In addition to changes to the laws of the game, decision training becomes even more important to retain consistency in the decision-making process of

referees. Published reports indicate that perceptual-cognitive training can improve decision making in referees (Catteeuw et al., 2010; Kittel et al., 2019; D. R. Mascarenhas, Collins, Mortimer et al., 2005; Schweizer et al., 2011). A correct and consistent initial decision is crucial because not all decisions can be reviewed, even though they can have an impact on the game dynamics. This suggestion highlights the need for referees to undertake appropriate training. Given the extra economic costs of using technology (Borooah, 2016), one could argue that more investment is needed in talent identification and training programmes with referees.

A second aim of this study was to investigate the number and duration of interventions by the VAR. One of the arguments against the use of the VAR is that the natural flow of a game may be negatively impacted by video technology (Ryall, 2012; Svantesson, 2014). Since a referee team in association football makes approximately 200–250 foul/no foul decisions per game (Helsen & Bultynck, 2004), the VAR protocol is restricted to match-changing incidents. Our data showed that on average only 4.4 checks were required per match. The median time duration taken when referring to the VAR was 22 seconds.

In more than 70% of all matches, the checks of the match-changing incidents did not lead to a review. Most VAR checks can occur in the background, thus having a minimal impact on the course of the game. In matches with a review, the median duration for an on-field review was 62.0 s and 15.0 s for a VAR-only review. A review is always preceded by a check and every time a review takes place, the referee has to indicate this by outlining a TV signal and the game has to be interrupted. A recent study showed that there was a slight increase in the total playing time and a simultaneous decrease in the effective playing time of matches with a review compared to matches without a review in the Spanish football league (Errekagorri et al., 2020). It is known that the effective playing time during a match (i.e., the total time during the match that the ball is in play) is 52–56 minutes (Castellano et al., 2011; Linke et al., 2018; Siegle & Lames, 2012). More than 30 minutes of playing time is lost due to other factors, such as free kicks, throw-ins and goal kicks (Augste & Cordes, 2016). When the ball is out of play, players actively position themselves for the next phase of play and the players and fans still have the feeling that the game is continuing. The duration of a review is relatively low in relation to these time losses. During an on-field review, however, the players, coaches and fans have to wait for a decision to be made while they often may not know what decision is being reviewed. National associations are therefore looking for innovative ways to communicate information relating to individual reviews to the TV production team, players and spectators in the stadium.

In other sports, like basketball and American football, a clock stop is used to compensate for the time losses. To reduce time-wasting and to speed up the game, the IFAB and FIFA have discussed the idea of stopping the clock every time the ball is out of play and adopt an effective playing time of 60 minutes (The International Football Association Board, 2017, "Increasing playing time", Chapter 2). Another viable option might be to adopt a stricter calculation of additional time to compensate for the time loss due to VAR or other interruptions.

An analysis of the decision changes showed that there were more penalties and red cards and fewer goals due to the interventions of the VAR. In future, researchers should investigate the (indirect) effects of the VAR and long-term changes on the game of association football. The presence of the VAR could have a deterrent effect that is not directly visible; it could eventually be that players act with more caution which will lead to fewer offences being committed and more goals being scored. A recent study showed that the number of fouls and yellow cards decreased after the implementation of the VAR in Germany and Italy (Carlos et al., 2019). The presence of the VAR appears to affect the (assistant) referee's performance outside of those decisions where VAR can be used. Further research is needed to clarify the effects of the VAR on the game and the behaviour of players and referees. Moreover, several factors such as the level of the match/competition/referee team and the score-line of the match could have an impact on the use and efficacy of VAR.

In sum, we showed that the predictive odds of making the correct decision were significantly higher when a decision was taken with the use of the VAR, compared to the initial decision made without the VAR. It is not possible to achieve 100% accuracy since there remains a degree of error in human perception and judgement, presenting a window for interpretation by the referees. In this respect, we highlighted the importance of a correct and consistent initial decision and the continuation, and even an expansion, of referee training programmes. With high initial decision accuracy, the number of VAR interventions and accompanying time losses are limited. It is important to continue applying video technology in a well-thought-out manner since it has the ability to undermine the referee team's credibility and change the character of the football game (Collins, 2010; Kolbinger & Lames, 2017). Fundamental to this latter aspiration is the need for more research in order to ensure that evidence-based practice is central to improvements in referee training and the use of new technologies.

Acknowledgments

The authors are very grateful to the IFAB for their cooperation and financial support in this research project.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by the The IFAB.

ORCID

Jochim Spitz http://orcid.org/0000-0002-2334-8734
Werner F. Helsen http://orcid.org/0000-0001-7066-8127

Data availability

The dataset supporting the conclusions of this article is available in an online repository (https://osf.io/c3agk).

References

Augste, C., & Cordes, O. (2016). Game stoppages as a tactical means in soccer–a comparison of the FIFA World Cups 2006 and 2014. *International Journal of Performance Analysis in Sport*, *16*(3), 1053–1064. https://doi.org/10.1080/24748668.2016.11868947

Balmer, N. J., Nevill, A. M., Lane, A. M., Ward, P., Williams, A. M., & Fairclough, S. H. (2007). Influence of crowd noise on soccer refereeing consistency in soccer. *Journal of Sport Behavior*, *30*(2), 130–145.

Borooah, V. K. (2016). Upstairs and downstairs the imperfections of cricket's decision review system. *Journal of Sports Economics*, *17*(1), 64–85. https://doi.org/10.1177/1527002513511973

Boyko, R. H., Boyko, A. R., & Boyko, M. G. (2007). Referee bias contributes to home advantage in English Premiership football. *Journal of Sports Sciences*, *25*(11), 1185–1194. https://doi.org/10.1080/02640410601038576

Carboch, J., Vejvodova, K., & Suss, V. (2016). Analysis of errors made by line umpires on ATP tournaments. *International Journal of Performance Analysis in Sport*, *16*(1), 264–275. https://doi.org/10.1080/24748668.2016.11868885

Carlos, L. P., Ezequiel, R., & Anton, K. (2019). How does video assistant referee (VAR) modify the game in elite soccer? *International Journal of Performance Analysis in Sport*, *19*(4), 646–653. https://doi.org/10.1080/24748668.2019.1646521

Caruso, E. M., Burns, Z. C., & Converse, B. A. (2016). Slow motion increases perceived intent. *Proceedings of the National Academy of Sciences of the USA*, *113*(33), 9250–9255. https://doi.org/10.1073/pnas.1603865113

Castellano, J., Blanco-Villaseñor, A., & Alvarez, D. (2011). Contextual variables and time-motion analysis in soccer. *International Journal of Sports Medicine*, *32*(6), 415–421. https://doi.org/10.1055/s-0031-1271771

Catteeuw, P., Gilis, B., Wagemans, J., & Helsen, W. (2010). Perceptual-cognitive skills in offside decision making: Expertise and training effects. *Journal of Sport and Exercise Psychology*, *32*(6), 828–844. https://doi.org/10.1123/jsep.32.6.828

Catteeuw, P., Helsen, W. F., Gilis, B., & Wagemans, J. (2009). Decision-making skills, role specificity, and deliberate practice in association football refereeing. *Journal of Sports Sciences*, *27*(11), 1125–1136. https://doi.org/10.1080/02640410903079179

Collins, H. (2010). The philosophy of umpiring and the introduction of decision-aid technology. *Journal of the Philosophy of Sport*, *37*(2), 135–146. https://doi.org/10.1080/00948705.2010.9714772

Coutts, A. J. (2016). Working fast and working slow: The benefits of embedding research in high performance sport. *International Journal of Sports Physiology and Performance*, *11*(1), 1–2. https://doi.org/10.1123/IJSPP.2015-0781

Errekagorri, I., Castellano, J., Echeazarra, I., & Lago-Peñas, C. (2020). The effects of the video assistant referee system (VAR) on the playing time, technical-tactical and physical performance in elite soccer. *International Journal of Performance Analysis in Sport*, Advanced online publication. https://doi.org/10.1080/24748668.2020.1788350

Fédération Internationale de Football Association. (2017). *Laws of the Game*. FIFA.

Fédération Internationale de Football Association. (2018). *Laws of the Game*. FIFA.

Fédération Internationale de Football Association. (2019). *Laws of the Game*. FIFA.

Gilis, B., Weston, M., Helsen, W. F., Junge, A., & Dvorak, J. (2006). Interpretation and application of the laws of the game in football incidents leading to player injuries. *International Journal of Sport Psychology*, *37*(2/3), 121.

Helsen, W., MacMahon, C., & Spitz, J. (2019). Decision making in match officials and judges. In A. M. Williams & R. C. Jackson (Eds.), *Anticipation and decision making in sport* (pp. 250–266). Routledge.

Helsen, W., & Bultynck, J. B. (2004). Physical and perceptual-cognitive demands of top-class refereeing in association football. *Journal of Sports Sciences*, *22*(2), 179–189. https://doi.org/10.1080/02640410310001641502

Johnson, J. G., & Raab, M. (2003). Take the first: Option-generation and resulting choices. *Organizational Behavior and Human Decision Processes*, *91*(2), 215–229. https://doi.org/10.1016/S0749-5978(03)00027-X

Jones, M. V., Paull, G. C., & Erskine, J. (2002). The impact of a team's aggressive reputation on the decisions of association football referees. *Journal of Sports Sciences*, *20*(12), 991–1000. https://doi.org/10.1080/026404102321011751

Kittel, A., Elsworthy, N., & Spittle, M. (2019). Incorporating perceptual decision-making training into high-intensity interval training for Australian football umpires. *Journal of Sports Sciences*, *37*(1), 29–35. https://doi.org/10.1080/02640414.2018.1480257

Kolbinger, O., & Lames, M. (2017). Scientific approaches to technological officiating aids in game sports. *Current Issues in Sport Science*, *2*(1), 1–10. https://doi.org/10.15203/CISS_2017.001

Leveaux, R. (2010). Facilitating referee's decision making in sport via the application of technology. *Communications of the IBIMA*, *2010*(2010), 1–8. https://doi.org/10.5171/2010.545333

Linke, D., Link, D., Weber, H., & Lames, M. (2018). Decline in match running performance in football is affected by an increase in game interruptions. *Journal of Sports Science & Medicine*, *17*(4), 662.

Mascarenhas, D. R., Button, C., O'Hare, D., & Dicks, M. (2009). Physical performance and decision making in association football referees: A naturalistic study. *The Open Sports Sciences Journal*, *2*(1), 1–9. https://doi.org/10.2174/1875399X00902010001

Mascarenhas, D. R., Collins, D., & Mortimer, P. (2002). The art of reason versus the exactness of science in elite refereeing: Comments on Plessner and Betsch (2001). *Journal of Sport and Exercise Psychology*, *24*(3), 328–333. https://doi.org/10.1123/jsep.24.3.328

Mascarenhas, D. R., Collins, D., & Mortimer, P. (2005). Elite refereeing performance: Developing a model for sport science support. *The Sport Psychologist*, *19*(4), 364–379. https://doi.org/10.1123/tsp.19.4.364

Mascarenhas, D. R., Collins, D., Mortimer, P. W., & Morris, B. (2005). Training accurate and coherent decision making in rugby union referees. *The Sport Psychologist*, *19*(2), 131–147. https://doi.org/10.1123/tsp.19.2.131

Myint, H., Wong, P., Dooley, L., & Hopgood, A. (2015, May). *Tracking a table tennis ball for umpiring purposes* [Paper presentation]. 14th IAPR international conference on machine vision applications, Tokyo.

Nevill, A. M., Balmer, N. J., & Williams, A. M. (2002). The influence of crowd noise and experience upon refereeing decisions in football. *Psychology of Sport and Exercise*, *3*(4), 261–272. https://doi.org/10.1016/S1469-0292(01)00033-4

Nlandu, T. (2012). The fallacies of the assumptions behind the arguments for goal-line technology in soccer. *Sport, Ethics and Philosophy*, *6*(4), 451–466. https://doi.org/10.1080/17511321.2012.745896

Pizzera, A., Marrable, J., & Raab, M. (2016). How effective is a video review system in soccer? *Research Quarterly for Exercise and Sport*, *87*(S1), S83.

Plessner, H., & Betsch, T. (2001). Sequential effects in important referee decisions: The case of penalties in soccer. *Journal of Sport and Exercise Psychology*, *23*(3), 254–259. https://doi.org/10.1123/jsep.23.3.254

Plessner, H., & Haar, T. (2006). Sports performance judgments from a social cognitive perspective. *Psychology of Sport and Exercise*, *7*(6), 555–575. https://doi.org/10.1016/j.psychsport.2006.03.007

Royce, R. (2012). Refereeing and technology–reflections on Collins' proposals. *Journal of the Philosophy of Sport*, *39*(1), 53–64. https://doi.org/10.1080/00948705.2012.675066

Ryall, E. (2012). Are there any good arguments against goal-line technology? *Sport, Ethics & Philosophy*, *6*(4), 439–450. https://doi.org/10.1080/17511321.2012.737010

Schweizer, G., Plessner, H., Kahlert, D., & Brand, R. (2011). A video-based training method for improving soccer referees' intuitive decision-making skills. *Journal of Applied Sport Psychology*, *23*(4), 429–442. https://doi.org/10.1080/10413200.2011.555346

Siegle, M., & Lames, M. (2012). Game interruptions in elite soccer. *Journal of Sports Sciences*, *30*(7), 619–624. https://doi.org/10.1080/02640414.2012.667877

Spitz, J., Moors, P., Wagemans, J., & Helsen, W. F. (2018). The impact of video speed on the decision-making process of sports officials. *Cognitive*

Research: Principles and Implications, *3*((16)), 1–10. https://doi.org/10.1186/s41235-017-0085-0

Spitz, J., Put, K., Wagemans, J., Williams, A. M., & Helsen, W. F. (2016). Visual search behaviors of association football referees during assessment of foul play situations. *Cognitive Research: Principles and Implications*, *1*(1), 12. https://doi.org/10.1186/s41235-016-0013-8

Spitz, J., Put, K., Wagemans, J., Williams, A. M., & Helsen, W. F. (2017). Does slow motion impact on the perception of foul play in football? *European Journal of Sport Science*, *17*(6), 748–756. https://doi.org/10.1080/17461391.2017.1304580

Svantesson, D. J. B. (2014). Could technology resurrect the dignity of the FIFA World Cup refereeing? *Computer Law & Security Review*, *30*(5), 569–573. https://doi.org/10.1016/j.clsr.2014.07.004

The International Football Association Board. (2017). *Play fair*. IFAB. https://play-fair.com

Unkelbach, C., & Memmert, D. (2008). Game management, context effects, and calibration: The case of yellow cards in soccer. *Journal of Sport and Exercise Psychology*, *30*(1), 95–109. https://doi.org/10.1123/jsep.30.1.95

Unkelbach, C., & Memmert, D. (2010). Crowd noise as a cue in referee decisions contributes to the home advantage. *Journal of Sport and Exercise Psychology*, *32*(4), 483–498. https://doi.org/10.1123/jsep.32.4.483

Van Quaquebeke, N., & Giessner, S. R. (2010). How embodied cognitions affect judgments: Height-related attribution bias in football foul calls. *Journal of Sport and Exercise Psychology*, *32*(1), 3–22. https://doi.org/10.1123/jsep.32.1.3