Journal of Sports Sciences
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713721847

Interpretation and application of the offside law by expert assistant referees: Perception of spatial positions in complex dynamic events on and off the field

Bart Gilis[a]; Werner Helsen[a]; Peter Catteeuw[a]; Evelien Van Roie[a]; Johan Wagemans[b]
[a] Department of Biomedical Kinesiology, [b] Department of Psychology, Katholieke Universiteit Leuven, Leuven, Belgium

To cite this Article Gilis, Bart , Helsen, Werner , Catteeuw, Peter , Van Roie, Evelien and Wagemans, Johan(2009) 'Interpretation and application of the offside law by expert assistant referees: Perception of spatial positions in complex dynamic events on and off the field', Journal of Sports Sciences, 27: 6, 551 — 563

To link to this Article: DOI: 10.1080/02640410802702178

URL: http://dx.doi.org/10.1080/02640410802702178

Journal of Sports Sciences, April 2009; 27(6): 551–563

Routledge
Taylor & Francis Group

Interpretation and application of the offside law by expert assistant referees: Perception of spatial positions in complex dynamic events on and off the field

BART GILIS[1], WERNER HELSEN[1], PETER CATTEEUW[1], EVELIEN VAN ROIE[1], & JOHAN WAGEMANS[2]

[1]*Department of Biomedical Kinesiology and* [2]*Department of Psychology, Katholieke Universiteit Leuven, Leuven, Belgium*

(Accepted 18 December 2008)

Abstract

Baldo et al. (2002) and Helsen et al. (2006) considered the flash-lag effect to explain errors made by assistant referees when judging offside in association football. The main aims of the present study were as follows: (1) to determine whether the flash-lag effect emerges in offside situations on the field of play or off the field when presented as computer animations or as video footage of real-life matches; (2) to examine offside decision-making errors in two standards of assistant referee – international FIFA and Belgian national referees. The results support the flash-lag hypothesis in several ways. First, both the FIFA and Belgian assistant referees were more likely to make errors by raising their flag when they had to assess offside situations on the field of play and when presented as three-a-side computer animations. Second, more flag errors were made when the defender moved in the opposite direction to that of the attacker. Third, the strategy of raising the flag in case of doubt was not observed when an interpretation of the offside law had to be made about the involvement of play of an attacker. Future research is needed to examine the extent to which on- and off-the-field training sessions can be used as training tools to improve offside decision-making.

Keywords: *Decision-making skills, expertise, flash-lag effect*

Introduction

In association football (referred to as soccer in North America), the responsibility of the match officials is to ensure that the laws of the game are properly respected by players and coaches. They therefore require the knowledge and the decision-making skills to interpret the laws of the game in a consistent and uniform way. In addition, they need to be physically fit to be able to move into the best possible position so as to make the most appropriate decision (for reviews, see Castagna, Abt, & D'Ottavio, 2007; Reilly & Gregson, 2006). They also need excellent perceptual skills to analyse every situation in a split-second from the start to the end of the game (Gilis, Weston, Helsen, Junge & Dvorak, 2006; Helsen, Gilis, & Weston, 2006; Mascarenhas, O'Hare, & Plessner, 2006). To improve the standard of refereeing, training sessions and the instruction of match officials need to focus on these different aspects (Reilly & Gregson, 2006).

In contrast to the considerable number of publications that have focused on the acquisition of expert perception and performance of athletes in various sports, such as basketball (Starkes, Allard, Lindley, & O'Reilly, 1994), field hockey (Helsen, Starkes, & Hodges, 1998; Starkes, 1987), and football (Helsen & Starkes, 1999; Helsen et al., 1998; Williams & Davids, 1995), few publications have considered these issues in match officials. Recently, Gilis and colleagues (Gilis, Helsen, Catteeuw, & Wagemans, 2008) found clear differences between assistant referees of the Fédération Internationale de Football Association (FIFA) and Belgian national-standard assistant referees in offside decision-making skills and recall accuracy in complex dynamic events using computer animations. Furthermore, it has also been shown that although coaches, players, and referees in football share a common knowledge base, referees outperformed coaches and players in a typical decision-making task in which the laws of the game had to be applied

Correspondence: W. Helsen, Department of Biomedical Kinesiology, Katholieke Universiteit Leuven, B-3000 Leuven (Heverlee), Belgium.
E-mail: werner.helsen@faber.kuleuven.be

ISSN 0264-0414 print/ISSN 1466-447X online © 2009 Taylor & Francis
DOI: 10.1080/02640410802702178

(Gilis et al., 2006; MacMahon, Helsen, Starkes, & Weston, 2006). Based on this research, it was concluded that there are role-specific skills within sports.

When experts in various domains exhibit these specific skills in public, their behaviour looks so effortless and people are tempted to attribute it to special talents. However, while hereditary factors are likely to play a role in shaping an individual's response to practice and training, skills are highly modifiable and adaptable to training. Therefore, every expert needs to practise for many hours to develop and refine these skills. In this regard, previous studies have indicated that a minimum of 10,000 h or 10 years of deliberate practice is required to attain expert performance in musicians (Ericsson, Krampe, & Tesch-Römer, 1993). Helsen et al. (1998) were the first to provide empirical evidence for this in team sports such as field hockey and football.

In addition to how naturally superior performance can be captured in the laboratory (Ericsson & Williams, 2007), a related challenge deals with the way in which training sessions on and off the field can be designed, implemented, and evaluated in the best possible way to improve perceptual-cognitive skills during match-play. For football refereeing, some studies have evaluated the effects of high-intensity and interval-based training sessions and ways to improve referees' physical capability and match performances (Krustrup & Bangsbo, 2001; Weston, Helsen, MacMahon, & Kirkendall, 2004). However, we are not aware of any studies that have focused on the acquisition of perceptual-cognitive skills in football referees and assistant referees.

An important issue concerns the extent to which training sessions contribute to improved perceptual-cognitive skills, such as anticipation and decision making (for a review, see Williams & Hodges, 2005). In this regard, there is evidence that the acquisition of perceptual-cognitive skills can be mediated through appropriate interventions. Several studies have shown that expert performers do not possess superior visual hardware to novice athletes (Helsen & Starkes, 1999). Generalized visual training programmes also do not improve vision or motor performance in sports (Abernethy & Wood, 2001). Instead, skilled performers appear to be able selectively to attend to, recognize, analyse, and interpret domain-specific visual information more effectively than their less skilled counterparts (for reviews, see Williams, 2000; Williams & Grant, 1999).

Although there have been some field-based training studies, the typical approach has involved video-based simulation. For example, Abernethy and Farrow (2002) revealed that schoolboy tennis players, who had access to training that required them to predict serve speed and direction while viewing temporally occluded video footage of the returns-of-serve, improved their prediction accuracy after the training intervention. Thus, improvements in perceptual skills did transfer to and improve task performance. In line with these findings, Williams and colleagues (Williams, Ward, & Chapman, 2002) concluded that anticipatory performance on a field hockey penalty flick can be improved by following a structured perceptual training programme using video simulation, instruction, and feedback. Again, a transfer from the laboratory to the field setting was apparent.

In association football, players, referees, and assistant referees should be able to refine their perceptual skills to achieve expert performance. Previous research with respect to the perceptual-cognitive demands of top-class refereeing showed that an expert official makes three to four technical decisions per minute throughout a match (Helsen & Bultynck, 2004). If, however, we also consider that a referee changes speed and direction during a game every 4 s (Krustrup & Bangsbo, 2001) to monitor the game from the most appropriate viewing angle, then it becomes clear that expert perception and performance complement each other. For assistant referees, a mean of 5.3 offside situations resulting in a flag signal per match was reported for the 2002 World Cup in Japan and Korea (Helsen et al., 2006). Although it is yet to be determined how many offside situations there are in which assistant referees do not raise the flag but let the game continue, assistant referees also show a change in speed and direction every 6 s to stay in line with the second last defender (Krustrup, Mohr, & Bangsbo, 2002). Again, this perceptual-cognitive involvement is apparent throughout a match to make correct offside and onside calls. While the number of publications examining decision-making processes in match officials in general is limited, this is even more so for the case of assistant referees and the typical offside decisions to which they are exposed.

At face value, the offside law is straightforward:

> a player is in an offside position if he is nearer to the opponents' goal line than both the ball and the second last opponent. A player in an offside position is only penalized if, at the moment the ball is touched or is played by one of his team, he is, in the opinion of the referee, involved in active play by interfering with play, or interfering with an opponent, or gaining an advantage by being in that position. "Nearer to his opponents' goal line" means that any part of his head, body or feet is nearer to his opponents' goal line than both the

ball and the second last opponent. The arms are not included in this definition. (FIFA, 2006)

Helsen et al. (2006) reported an error rate of 26.2% for top-class FIFA assistant referees in assessing offsides during the 2002 FIFA World Cup. In addition, a bias was found towards flag errors (the assistant referee raises his flag while the attacker is in an onside position at the moment the ball is played) compared with non-flag errors (the assistant referee fails to raise his flag while the attacker is in an offside position).

Accordingly, it is important to gain a better understanding of these erroneous judgements. Baldo and colleagues (Baldo, Ranvaud, & Morya, 2002) suggested the flash-lag hypothesis for explaining the underlying mechanisms of errors made in judging offside. The flash-lag effect refers to a perceptual illusion in which a moving object is perceived as spatially leading its real position at an instant defined by a time marker (usually a briefly flashed stimulus) (Nijhawan, 1994). In the case of offsides in football, the attacker receiving the ball is a moving stimulus, often progressing at high speed. The moment at which the ball is passed can be interpreted as a flash that clearly marks the instant at which the position of the attacking player has to be judged by the assistant referee. As a result, the attacker receiving the ball, often moving forward, is perceived ahead of his actual position due to the flash-lag effect. This leads to the expectation of an overall bias towards flag errors compared with non-flag errors. Recently, Gilis et al. (2008) found evidence for the flash-lag effect in complex dynamic events using computer animations. Figure 1 provides an illustration of the flash-lag effect in offside situations.

In the present study, the specific offside decision-making skills of expert assistant referees were first evaluated by focusing on the first part of the offside

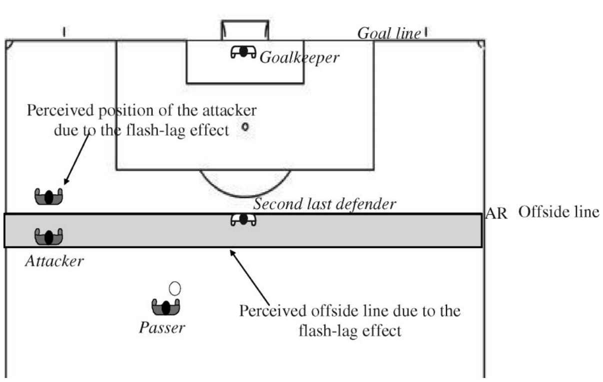

Figure 1. Illustration of the flash-lag effect. The attacking player, who is in a forward motion, is perceived ahead of his true position. This results in an increased probability of flag errors when the attacking player is in the shaded area. AR = assistant referee.

law, how assistant referees perceive the relative positions of the attackers and defenders at the moment the ball is played. The second part of the offside law – whether the attacking player was involved in active play, interfering with play, interfering with an opponent, or gaining an advantage by being in that position – was examined using video footage of real matches. For these purposes, we used data collected during a FIFA workshop to select the best assistant referees for the 2006 FIFA World Cup in Germany.

To evaluate the specific offside decision-making skills, we used data from both on- and off-the-field offside tests. Youth football players attending a football academy in Germany simulated offside situations for the on-the-field test. For the off-the-field offside test, a well-controlled data set of offside situations was presented using computer animations. To test the interpretation of the offside law, video footage of real matches was used in which assistant referees had to decide whether the attacker was involved in active play, interfered with an opponent or gained an advantage by being in an offside position.

More specifically, the first aim of this study was to examine types of error made by FIFA assistant referees in an on-the-field offside test. The error rate was anticipated to be higher when an attacker was located closer to the offside line. However, in line with the forward displacements due to the flash-lag effect, the forward moving attacker would be perceived ahead of his actual position at the moment the ball is played. This would result in an overall bias towards flag than non-flag errors. Based on the forward displacements associated with the flash-lag effect, we assumed that more errors would be made when the defender moved in the opposite direction to that of the attacker compared with offsides where the defender was standing still at the moment the ball was played, because the defender's movement will also give rise to its own flash-lag effect (but in the opposite direction). Furthermore, the position of the assistant referee relative to the offside line was evaluated in line with the correctness of his decision. Based on the optical error hypothesis, as proposed by Oudejans et al. (2000), it was anticipated that the position of the assistant referee would influence the correctness of his decision.

Our second aim was to investigate the types of error made by FIFA assistant referees when judging offside situations presented as computer animations. In line with previous research (Gilis et al., 2008; Helsen et al., 2006), we hypothesized that the flash-lag effect would have an impact on the nature of the decision in the same way as in the on-the-field test. To assess this in more detail, the action of the defender at the moment the ball was passed was experimentally manipulated. In the first condition, the defender stood completely still while the attacker moved towards the goal line. In the second condition, the defender moved in the opposite direction of the receiving attacker. It was anticipated that the latter case would lead to more errors than the static condition, because not only will the attacker be seen ahead of his position, but also the defender, although in the opposite direction. A second manipulation was the speed of the offside situations, with a frame rate of 6 or 8 frames per second. Animations played at 8 frames per second were hypothesized to be more difficult than those played at a lower speed of 6 frames per second. This assumption is based on research by Finke and colleagues (Finke, Freyd, & Shyi, 1986), who stated that memory distortions as a cause of forward displacements are sensitive to changes in the implied speed. Also, for continuously moving targets, larger forward displacements were reported when speed increased (Hubbard & Bharucha, 1988).

In addition, the same computer animations were shown to a group of 36 Belgian assistant referees in an attempt to discriminate between different standards of expertise. We predicted that international FIFA assistant referees would outperform their Belgian national-standard colleagues in the number of correct decisions. Gilis et al. (2008) previously found differences in response accuracy between assistant referees of different standards. In this study, we wanted to determine whether these findings could be replicated.

The third aim was to investigate the declarative knowledge of the offside law. FIFA and Belgian assistant referees had to interpret the offside rule in difficult situations. We hypothesized that the FIFA experts would make fewer errors in this specific knowledge test than the Belgian assistant referees. We wondered whether there would be a tendency for assistant referees to raise their flag in this type of offside situation.

Our final aim was to compare the results on the laboratory offside tests of the top-class assistant referees with the on-the-field offside test. This kind of investigation is important to develop future training and selection procedures of assistant referees.

Methods

Participants

The participants were 70 FIFA assistant referees (mean age 40.0 years, $s = 4.2$, range $= 29.4–45.1$) attending a one-week FIFA workshop for prospective assistant referees in Frankfurt, Germany, 7 weeks before the start of the 2006 FIFA World Cup in Germany. These assistant referees, selected from around the world, were all potential candidates for nomination to the 2006 FIFA World Cup. Their mean experience as an international FIFA assistant referee was 6.7 years ($s = 3.2$, range $= 1–13$).

During the workshop, the assistant referees undertook several tests. First, they had to pass a physical fitness test that consisted of six repeated 40-m sprints and 20 interval runs (150 m in 30 s alternated with 50 m walking for 40 s). In terms of offside decision-making tests, the assistant referees were exposed to three formats. The first format was an on-the-field offside test with typical one-to-one offside situations. The second was an off-the-field test that consisted of computer animations that represented typical offside challenges. And the third was an off-the-field offside decision-making test based on video footage of real matches. Selection of the assistant referees for the 2006 FIFA World Cup was based on the results of these three tests, combined with match reports and refereeing performances during national and international matches.

To validate the laboratory-based offside decision-making tests and to examine the extent to which these tests could discriminate between assistant referees of different levels of expertise, a second group of participants consisting of 36 Belgian assistant referees (mean age $= 38.5$ years, $s = 4.4$, range $= 24.2–44.4$) was tested. These assistant referees had experience in the first and second division of the Belgian national football league for a mean of 4.3 years ($s = 3.1$, range $= 1–11$). The tests were carried out during a regular course in Leuven, Belgium. Following a brief explanation of the nature of the investigation, written consent was obtained from the FIFA and the Belgian referees' committees. The study was designed and conducted in accordance with the ethical standards laid down in the 1964 Declaration of Helsinki and approved by the Committee for Ethical Considerations in Human Experimentation of the Faculty of Kinesiology and Rehabilitation Sciences, Katholieke Universiteit Leuven.

On-the-field offside test

Task and apparatus. An on-the-field offside decision-making test was designed with a fixed-test protocol in which one player who passed the ball, one attacker, and one defender were included. For this field experiment, a team of 15- to 16-year-old football players, all attending a football academy in Germany, were instructed to simulate offside situations on a regular field with natural grass. The ball was always kicked from a steady position and the initial positions of the attacker and defender were kept constant.

The attacker, running at a moderate speed of 12–14 km · h^{-1}, always moved in a forward direction. The movement of the defender was experimentally manipulated to have two conditions. In the first condition, the defender was standing still at the moment the pass was made, whereas in the second, the defender moved in the opposite direction of the attacker at a similar moderate speed. As during real matches, the assistant referees were instructed to be in line with the second last defender. In Figure 2, the organization of the on-the-field offside test is presented with the exact distances.

Before the test, clear instructions were given to the players about the protocol and the objectives of this offside test. In addition, a 30-min practice session was organized to accustom the players to the test. In this respect, appropriate feedback was given after every trial.

Procedure. The assistant referee was positioned on the touchline and had to make a decision by raising his flag when the attacker was in an offside position or keeping his flag down when the attacker was in an onside position at the moment the pass was made. Each assistant referee was exposed to 30 offside situations. The test was carried out in groups of eight assistant referees, in such a way that the assistant referees alternated and followed one another after the first trial before assessing the second offside situation. First, the test with the static defender ($n = 15$) was undertaken and, one day later, the assistant referees were exposed to situations with a dynamic

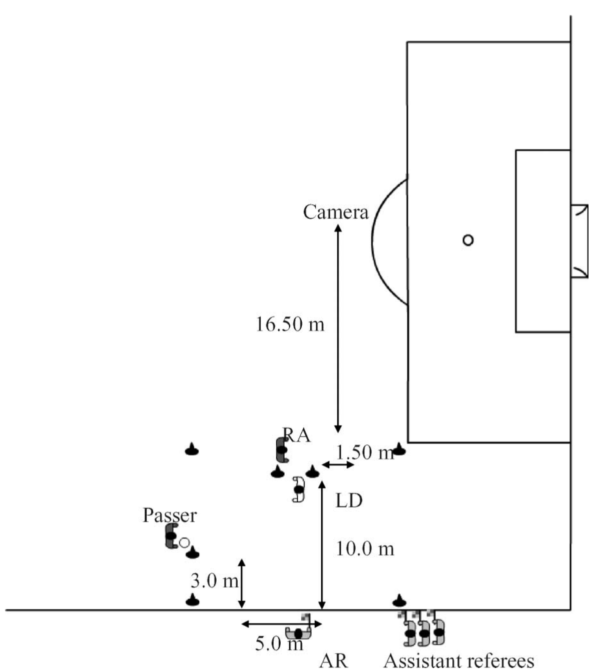

Figure 2. Overview of the on-the-field offside test with one passing attacker, one receiving attacker (RA), one second-last defender (LD), and the assistant referee (AR) on the touchline.

defender ($n = 15$). Every test session lasted 20 min. No feedback was provided to the assistant referees about their performance during or after the test session.

Data reduction. The on-the-field decision-making test was captured with a digital video camera (Canon, DV Camcorder MV4). On the video image, the actions of the passer, attacker, second last defender, and the assistant referee were all visible. First, the video footage was imported in the software program Final Cut Pro 5.1.4, which allowed us to stop and play back all the situations to examine the correctness of the assistant referee's judgements at the precise moment the ball was played. A distinction was made between a correct flag signal when the attacker was in an offside position and a correct non-flag signal when the attacker was in an onside position. In contrast, when the assistant referee made an incorrect decision, either he raised his flag even though the attacker was in an onside position, thus making a flag error, or he kept his flag down when the attacker was in an offside position, thus making a non-flag error. Second, using Final Cut Pro 5.1.4 we were able to freeze the video footage and take a picture of every offside situation at the moment the ball was played to analyse the distance between the attacker and the defender and to evaluate the position of the assistant referee relative to the offside line. Third, to correct the perspective bias of this picture, we used the photo editing program GIMP 2.2.13 (GNU Image Manipulation Program) to straighten the picture. A rectangular box with gridlines was placed on top of each picture and the four nodes of the box were placed on the reference points on the picture. After the transformation of the perspective, the distances at the bottom and the top of the picture were exactly the same. Then, a tailor-made software program was created in Delphi 6 Professional to measure the distances between the attacker and the second-last defender and between the assistant referee and the second-last defender. A marker was placed on both the defender and the attacker on the part of the body (except the arms) that was closest to the defender's goal line. The marker for the defender was equal to the offside line and was also used to evaluate the position of the assistant referee. The reference point for the assistant referee was chosen in the middle of his face. Together with the FIFA Head of Refereeing, we determined that the assistant referee was positioned on the offside line when he was exactly on the offside line (0 pixels or the part of the defender that was closest to the defender's goal line, mostly the foot of the defender) or not more than 35 cm behind the offside line (the body of the defender). Furthermore, the assistant referee was said to be positioned behind the offside line when he was more than 35 cm behind the line and ahead of the offside line when he was closer to the goal line than any part of the defender's body. The distances were measured in pixels, with one pixel corresponding to 8 mm. In just a few offside situations ($n=8$), the assistant referee was completely covered by the attacker and the defender so that his correct position could not be determined exactly. Therefore, these situations were not included in the final analysis.

Data analysis. Examination of the accuracy of the offside decisions as well as evaluation of the distances was done by two observers. To assess the objectivity, these two observers analysed the same data set of 40 situations, chosen at random. For the accuracy of the offside decisions, the inter-observer agreement was examined and expressed as a percentage. These two observers also analysed the distances between the attacker and the defender on the one hand and between the assistant referee and the offside line on the other. The intra- and inter-observer reliability of the two observers analysing the distances was tested by Wilcoxon matched-pairs signed-ranks test.

In line with previous work (Gilis et al., 2008), a chi-squared test was used to determine whether the ratio of correct versus incorrect decisions was uniformly distributed throughout the conditions with a static or dynamic defender, and also whether there was an overall bias towards flag errors compared with non-flag errors. In addition, error scores according to the distance between attacker and defender were calculated and differences were tested with a chi-square analysis. Furthermore, we used signal-detection theory (Macmillan & Creelman, 2005) to gain a better understanding of the performance of the FIFA assistant referees on the on-the-field offside test. The sensitivity index (d') was computed by comparing the hit rate (i.e. correct flags) to the false alarms (i.e. flag errors). In general, a high false alarm rate, resulting in a lower d' value, indicates less perceptual sensitivity. Specifically, a sensitivity index not different from zero implies that the assistant referees cannot distinguish between an offside and an onside position of the attacker. In addition, the response bias (c) was calculated to investigate the preference of the participants to flag or keep the flag down in doubtful situations. A negative c value implies that the assistant referees were more likely to flag, whereas a positive response bias indicates a preference for keeping the flag down. An alpha of 0.05 was adopted for all statistical tests.

In addition to the on-the-field test, the assistant referees were also exposed to two off-the-field offside decision-making tests. The first off-the-field test comprised offside situations presented as computer animations where an assessment had to be made of the players' positions at the moment the pass was made (i.e. the first part of the offside law). In the second, video footage of real matches was selected to assess the second part of the offside law: was the attacker involved in active play, interfering with play, interfering with an opponent, or gaining an advantage by being in that position.

Off-the-field offside test: Computer animations

Tasks and procedures. As in previous work (Gilis et al., 2008), the software program Macromedia Flash MX Professional 2004 (version 7.2) was used to create computer animations of offside situations. We began by drawing a standardized football pitch with dimensions as defined by the FIFA's (2006) laws of the game. Then attackers and defenders (6.8×20.9 pixels) were added, in red and yellow respectively. They could be moved in all directions and with various speeds. The three-a-side computer animations used in this study comprised three attackers, two defenders, and one goalkeeper. We created four typical patterns of play, all of which resulted in a potential offside situation. Within these four typical patterns, the position of the attacker relative to the offside line was manipulated experimentally. Specifically, four spatial positions were used, namely the attacker in position (1) 20 pixels behind the offside line, (2) 10 pixels behind the offside line, (3) exactly in line with the second-last defender (0 pixels), and (4) 10 pixels ahead of the offside line (Figure 3). Thus, only in the latter manipulation was the attacker in an offside position at the moment the ball was played towards him.

Second, we also experimentally manipulated the action of the defender. In the first condition, the defender stood completely still while the attacker was moving towards the goal line at the moment the pass was made. In the second condition, the defender moved in the opposite direction of the attacker. Third, the speed of the offside situations was manipulated. The computer animations were created at either 6 or 8 frames per second. For a better understanding of the value of a pixel, some references are given based on Figure 3: the dimension of the field of play was 605.0×289.8 pixels; the distance between the goal line and the halfway line was 302.5 pixels; and the distance from the goal line to the penalty box and to the goal area was 91.5 and 30.2 pixels respectively. In total, we created 64 three-a-side computer animations that were uniformly distributed across the various manipulations.

The start of every trial was indicated by a blue circle that appeared for 1.3 s around the player in possession of the ball to ensure that participants focused their attention on the task. The duration of one trial was 6 s. The trials were presented on a large

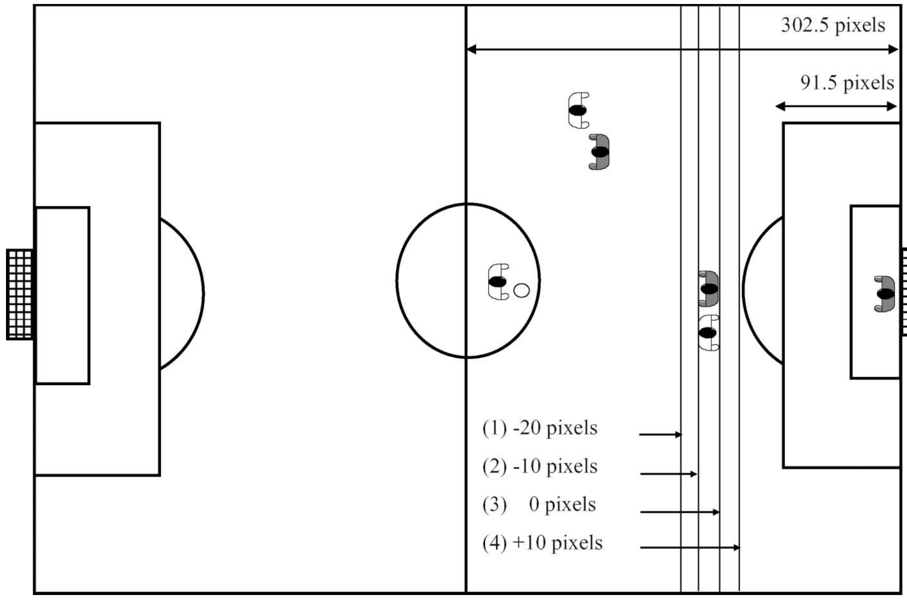

Figure 3. Illustration of the four potential positions of the attacker relative to the offside line with the attacker (1) 20 and (2) 10 pixels behind the offside line, (3) on the offside line (0 pixels) for the onside positions, and (4) with the attacker 10 pixels ahead of the offside line.

screen (2.60 × 3.60 m) and standardized instructions were given before the test. Also, five practice trials were given. For this laboratory test, the FIFA assistant referees were divided into three groups. All participants were seated in front of the screen and had an individual voting box with two alternatives, corresponding to "no offside" and "offside". The assistant referees had to decide whether the attacker was in an offside position at the precise moment the ball was played. The participants were given 10 s to press the appropriate button, after which the answers were automatically registered by a computer (Dell Laptop, Latitude D800). Similar computer animations and the same technique were used by Gilis et al. (2008) and shown to be successful in capturing specific offside decision-making skills.

Off-the-field offside test: Video footage of match-play

Twenty clips were selected from matches based on a database of FIFA and the Perception and Performance Laboratory of the Department of Biomedical Kinesiology, K.U. Leuven. This database consists of matches from various FIFA tournaments, including the 2002 FIFA World Cup in Korea and Japan, the 2005 FIFA Confederations Cup in Germany, and the 2005 FIFA Under-17 World Cup in Peru. The selection of offside situations was made together with the FIFA Refereeing Department and dealt with the second part of the offside law – namely, the assistant referee had to decide whether the attacker was involved in active play, interfered with play, interfered with an opponent, or gained an advantage by being in that position. First, the offside situation was presented, immediately followed by a replay in slow motion from a different viewing angle. Second, the assistant referees had to decide whether the attacker should be penalized for being offside. For every situation, they noted offside or no offside. Before testing started, the match situations were also presented to four experts of the FIFA Refereeing Department to obtain their views. These experts unanimously agreed on every match situation whether the flag needed to be raised or kept down. In total, there were 10 offside situations and 10 not offside situations.

Data analysis. These off-the-field offside decision-making tests were not only carried out by the prospective FIFA assistant referees, but also by a group of Belgian assistant referees from the Belgian national football league. First, the decisions of the assistant referees were compared with the correct answers, and then the overall correctness score or the response accuracy for each individual was calculated, expressed as a percentage. Second, the assumption of normality was checked using the Shapiro-Wilk *W*-test. To discriminate between the FIFA and the Belgian assistant referees, an unpaired *t*-test was used to compare the results of both groups on the computer animations and the real-life test, with the response accuracy of the assistant referees, expressed as a percentage, as the dependent variable. Third, a chi-squared test was used for the computer animations to examine if the ratio of correct versus incorrect decisions was uniformly distributed across the action of the defender (static vs. dynamic) and across the speed of the computer animations (6 vs. 8 frames per second). Also, we examined whether there was an overall bias towards flag errors in comparison with non-flag errors. Fourth, similar to the on-the-field offside test, signal-detection theory was used to measure the performance obtained for the computer animations. Additionally, to study the relationship between the laboratory tests and the on-the-field offside tests, Spearman rank correlation coefficients were calculated.

Results

On-the-field offside test

Intra- and inter-observer reliability. As two observers analysed the offside situations on the field of play, we examined the intra- and inter-observer reliability. First, using the same data set of offside situations ($n = 40$), these two observers assessed whether the attacker was in an onside or an offside position at the precise moment the pass was made and that this corresponded with the decision of the assistant referee. Inter-observer agreement of 97.5% was found between the two observers, which means that in only one of the situations was there no agreement. When one observer assessed this set of offside situations a second time, intra-observer agreement of 100% was found. Second, when these two observers estimated the distances between the attacker and the defender for the same data set, no significant differences were observed regarding intra-observer ($t_{40} = 286$, $P = 0.15$) and inter-observer reliability ($t_{40} = 262$, $P = 0.11$). There were also no differences between the two observers when they calculated the distances between the attacker and the defender ($t_{40} = 297$, $P = 0.19$) and when one of the observers re-calculated these distances ($t_{40} = 265$, $P = 0.13$).

Type of error. The overall error rate obtained for the on-the-field offside test was 25.9%. To examine types of error in more detail, we calculated the percentage of flag and non-flag errors. In total, 480 flag errors were made on 1739 occasions with the attacker in an onside position, resulting in an error rate of 27.6%. Second, there were 353 situations with the attacker in an offside position. The assistant referees were incorrect and kept their flag down in 61 of these situations (non-flag errors = 17.3%). As predicted by the flash-lag effect, there was indeed an overall bias towards flag errors in this on-the-field offside test ($\chi^2 = 24.83$, $P < 0.001$). The results also showed that more errors were made in the dynamic (319/1048; 30.4%) than in the static condition (222/1044; 21.3%) ($\chi^2 = 17.03$, $P < 0.001$). Finally, as anticipated, more flag errors (294/888; 33.1%) than non-flag errors (25/160; 15.6%) were made ($\chi^2 = 20.03$, $P < 0.001$) for situations in which the defender moved in the opposite direction to that of the attacker.

Oudejans et al. (2000, 2005) proposed the optical error hypothesis and observed that the position of the assistant referee is important to explain the type of error that is made. Therefore, to evaluate the types of error made by the assistant referees, the position relative to the offside line needs to be considered. First, the results showed that the assistant referees were further behind the offside line than considered appropriate (1155/2092; 55.2%), positioned on or just behind the offside line (717/2092; 34.2%), or ahead of the offside line (220/2092; 10.5%) ($\chi^2 = 627.66$, $P < 0.001$). More importantly, a bias towards flag errors compared with non-flag errors was identified for each position of the assistant referee relative to the offside line: behind ($\chi^2 = 6.51$, $P < 0.02$), on ($\chi^2 = 16.14$, $P < 0.001$) or ahead of ($\chi^2 = 11.81$, $P < 0.001$) the offside line (Table I). This observation is at odds with the predictions of the optical error hypothesis.

Table I. Number of correct non-flag signals (CNF), correct flag signals (CF), flag errors (FE), and non-flag errors (NFE) according to the position of the FIFA assistant referees (n = 70) – behind, on or ahead of the offside line – and according to the action of the defender (static or dynamic).

		Decision of FIFA assistant referees				
Defender's action	Position	CNF	CF	FE	NFE	Total
Static	Behind	387	31	83	9	510
	On the line	252	89	83	23	447
	Ahead	26	37	20	4	87
	Total	665 (78.1%)	157 (81.3%)	186 (21.9%)	36 (18.7%)	1044
Dynamic	Behind	432	46	161	6	645
	On the line	112	52	98	8	270
	Ahead	50	37	35	11	133
	Total	594 (66.9%)	135 (84.4%)	294 (33.1%)	25 (15.6%)	1048
Overall total		1259 (72.4%)	292 (82.7%)	480 (27.6%)	61 (17.3%)	2092

We also considered the distance between the attacker and the defender to gain a better understanding of the offside decision-making errors. In Figure 4, the continuous line represents the flag errors and the dotted line the non-flag errors. First, the results showed that for every distance between the attacker and the defender, more flag errors were made than non-flag errors ($P < 0.05$). When we examined the flag errors in more detail, an increase in the error rate was observed when the distance between the attacker and the defender decreased. In addition, if the distance between the attacker and the offside line was less than 0.5 m, the error score increased even more. In contrast to these findings, there was a levelling off in the percentage of non-flag errors when the distance between the attacker and the offside line was less than 0.5 m.

Using signal-detection theory, FIFA assistant referees were shown to discriminate between an offside and an onside position above chance ($d' = 1.54$, $P < 0.05$). The negative c value, however ($c = -0.19$), indicates that these top-class assistant referees were more likely to flag in doubtful and difficult situations, an observation that was also evidenced by the number of flag errors.

Off-the-field offside test: Computer animations

To determine whether the computer animations discriminated between expert and less experienced assistant referees, we compared the error rate of FIFA ($n = 68$) and Belgian assistant referees ($n = 36$). The results showed that the FIFA experts (73.7%) were more accurate in their decisions than the Belgian assistant referees (69.3%) ($t_{102} = 2.13$, $P = 0.036$). In addition, according to signal-detection theory, the d' values revealed that both the FIFA ($d' = 1.72$) and the Belgian ($d' = 1.65$) assistant referees could discriminate between offside and not offside above chance. No difference was found when

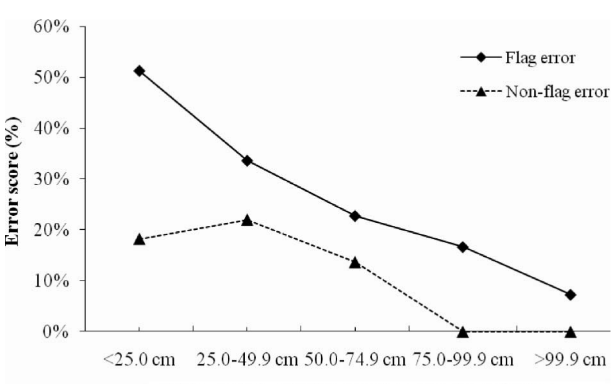

Figure 4. Error score (%) according to the distance between the attacker and the defender.

d' values for the FIFA and Belgian assistant referees were compared for situations with the static defender. However, situations with the dynamic defender revealed differences between FIFA ($d' = 1.89$) and Belgian ($d' = 1.40$) assistant referees ($P < 0.05$). As such, FIFA assistant referees were better at discriminating between offside and not offside in dynamic situations than their Belgian national-standard counterparts. Furthermore, the results showed a difference for the criterion (c) between the two standards ($P < 0.05$). Both values were negative, which means that the two groups prefer to flag in doubtful situations. However, the Belgian assistant referees ($c = -0.52$) were more willing to raise their flag than the FIFA assistant referees ($c = -0.37$).

Also in these computer animations, we wished to determine the number and types of errors. First, an overall bias towards flag errors (31.3%) compared with non-flag errors (11.3%) was found ($\chi^2 = 204.36$, $P < 0.001$). Furthermore, more errors were made by FIFA assistant referees in those situations with the attacker 10 pixels behind (-10 pixels) the offside line (257/1088; 23.6%) than in situations with the attacker 20 pixels behind (-20 pixels) (108/1088; 9.9%) or 10 pixels ahead ($+10$ pixels) of the offside line (123/1088; 11.3%). Also, the error rate was even higher when the attacker was on the offside line (0 pixels) (657/1088; 60.4%) ($\chi^2 = 687.28$, $P < 0.001$). Moreover, more errors were made when the defender was moving in the opposite direction to that of the attacker (647/2176; 29.7%) than when the defender was static (498/2176; 22.9%) ($\chi^2 = 19.40$, $P < .001$). More specifically, the number of flag errors was greater with a dynamic defender ($\chi^2 = 26.96$, $P < 0.001$). Finally, more errors were made on animations played at a speed of 8 frames per second (628/2176 or 28.9%) than animations at 6 frames per second (517/2176; 23.8%) ($\chi^2 = 10.76$, $P < 0.005$). Specifically, FIFA assistant referees made more flag errors when the animations were played at a speed of 8 frames per second ($\chi^2 = 9.40$, $P < 0.005$) (Table II).

Relationship between the on-the-field offside test and the computer animations

To compare the on-the-field and the off-the-field offside test, the degree of difficulty was equalized. Therefore, only those on-the-field situations were chosen where the attacker was less than 75 cm behind or ahead of the offside line. All together, each of the 68 assistant referees were exposed to at least 10 onside situations (mean = 16.2) with the attacker less than 75 cm behind the offside line. The results showed a relationship (Spearman rank correlation coefficient) for the mean of the scores in situations where the attacker was 10 pixels behind or on the

Table II. Number of correct and incorrect decisions of the FIFA assistant referees (n = 68) according to the position of the attacker relative to the offside line (−20, −10, 0, and +10 pixels), the action of the defender (static or dynamic), and the speed of the action (6 and 8 frames per second).

Position of attacker	Defender's action	Speed (frames · s^{-1})	Decision of FIFA assistant referees	
			Correct	Incorrect
−20 pixels	Static	6	266 / 97.8%	6 / 2.2%
		8	251 / 92.3%	21 / 7.7%
	Dynamic	6	236 / 86.8%	36 / 13.2%
		8	227 / 83.5%	45 / 16.5%
−10 pixels	Static	6	234 / 86.0%	38 / 14.0%
		8	206 / 75.7%	66 / 24.3%
	Dynamic	6	210 / 77.2%	62 / 22.8%
		8	181 / 66.5%	91 / 33.5%
0 pixels	Static	6	125 / 46.0%	147 / 54.0%
		8	122 / 44.9%	150 / 55.1%
	Dynamic	6	99 / 36.4%	173 / 63.6%
		8	85 / 31.3%	187 / 68.7%
+10 pixels	Static	6	243 / 89.3%	29 / 10.7%
		8	231 / 84.9%	41 / 15.1%
	Dynamic	6	246 / 90.4%	26 / 9.6%
		8	245 / 90.1%	27 / 9.9%
		Total	3207 / 73.7%	1145 / 26.3%

offside line in the computer animations and the on-the-field situations where the attacker was also within the limit of 75 cm behind the offside line ($\rho = 0.49$, $P < 0.05$). There was a relationship between the situations in which the attacker was 10 pixels ahead of the offside line in the computer animations and the on-the-field offside situations where the attacker was within 75 cm ahead of the offside line ($\rho = 0.34$, $P < 0.05$).

Off-the-field offside test: Video footage of matchplay

Results for the interpretation of the offside law revealed a difference in response accuracy between FIFA (81.8%, $s = 8.34$, range = 63.2–100) and Belgian (75.1%, $s = 7.29$, range = 63.2–89.5) assistant referees ($t_{104} = 4.06$, $P < 0.001$). The FIFA assistant referees ($d' = 1.837$) made more correct interpretations of the offside law than the Belgian assistant referees ($d' = 1.343$). Both the FIFA ($c = -0.107$) and the Belgian ($c = 0.054$) assistant referees were not biased to raise their flag or to keep their flag down in the case of doubtful situations.

Discussion

The offside law is one of the most debated laws in football. Helsen et al. (2006) reported an error rate of 26.2% during the 2002 FIFA World Cup in Korea and Japan over 64 matches. A similar error rate of 20% was observed by Oudejans et al. (2005) in four matches from the Dutch national league for situations where the attacker was less than 1 m away from the second-last defender. Assistant referees are advised to position themselves in line with second-last defender and to have an appropriate running technique, preferably adopting sideways movements so that their point of gaze is always focused within the field of play. In addition, they are advised to wait and see and to freeze the position of the attacker and the defender at the moment the ball was played and afterwards to judge the involvement in active play.

Previous research has suggested various explanations for incorrect judgements in offside situations. Sanabria et al. (1998) suggested that an assistant referee has to make a shift of gaze from the passer to the receiver, resulting in a time delay. Oudejans et al. (2005) referred to the speed of locomotion of the assistant referee. Specifically, more errors were made when an assistant referee was running or sprinting than when standing still, walking or jogging (Oudejans et al., 2005). Another possible determinant of the quality of offside decisions is the fact that the sprint performance of an assistant referee decreases towards the end of a game as a consequence of fatigue. Therefore, assistant referees might not be able to keep up with play, resulting in a less precise positioning with the offside line (Krustrup et al., 2002). Interestingly, however, Helsen et al. (2006) showed that the number of offside errors did not increase as the match progressed. Furthermore, Oudejans et al. (2000, 2005) suggested that the position of the assistant referee in relation to the offside line could have a marked impact on the offside judgement. As a result of an inappropriate position and corresponding angle of view, optical errors could be an important underlying mechanism of incorrect decisions.

From a perceptual point of view, Baldo et al. (2002) and Helsen et al. (2006) suggested the flash-lag effect to explain errors in judging offside during actual match-play. In line with these findings, Gilis et al. (2008) suggested the flash-lag effect to explain the bias towards flag errors in complex dynamic events presented as computer animations. The present study is the first to reveal the flash-lag effect in a real-life setting. A first important finding in this on-the-field test with one passer, one attacker, and one defender was the overall bias of the FIFA assistant referees towards flag errors (27.6%) compared with non-flag errors (17.3%). In addition, this bias was apparent for every position of the assistant referee relative to the offside line (i.e. behind, on or ahead of the offside line). Furthermore, in situations with a dynamic defender who was running in the opposite direction to the attacker, even more flag errors (33.1%) were made by the assistant referees. From these results, it became clear that the situations with the attacker just behind the offside line (51.2%) were more difficult to assess than the situations with the attacker just ahead of the line (18.2%). In the on-the-field test, therefore, the results revealed that the attacker was perceived ahead of his actual position at the moment the ball was played. When the defender moved in the opposite direction, the impact of the flash-lag effect was even stronger. In this test, the position of the assistant referee relative to the offside line was not a determinant of incorrect offside decision making, since the attacker and defender crossed each other shoulder to shoulder. Therefore, the distance between the defender and the attacker in the width of the field of play was minimal and the deviation of the assistant referee relative to the offside line was negligible.

Mascarenhas et al. (2006) suggested a signal-detection method to determine whether assistant referees are more motivated towards detecting offside and thus make more flag errors than non-flag errors. Signal-detection theory results found in the present study demonstrated the usefulness of this concept. Specifically, the theory revealed that top-class FIFA assistant referees could discriminate between an offside and an onside position of the attacker at the specific moment the pass is given. However, in case of doubt, the response bias showed that the assistant referees were more likely to raise their flag.

Another important result of the present study that is in line with the research of Gilis et al. (2008) is that offside situations presented as computer animations discriminated between FIFA and Belgian assistant referees. FIFA assistant referees were more accurate in assessing offside (73.7%) than the Belgian (69.3%) assistant referees. The results of the signal-detection analysis revealed also that FIFA assistant referees were more sensitive to discriminate between offside and not offside in dynamic situations. Although both groups were more likely to flag in case of doubt, Belgian assistant referees raised their flag even more than FIFA assistant referees. Thus, it appears that offside decision making in computer animations is task specific and that expertise can clearly be captured through this laboratory test.

While the on-the-field offside test and the offside test based on computer animations mainly investigated the skills of assistant referees to make a correct offside assessment purely based on the position of the attacker, another aim of this study was to examine how assistant referees dealt with the second part of the offside law, namely whether the attacker in an offside position is involved in active play. This was the first time that this part of the offside law was tested. The results showed a difference in response accuracy of FIFA assistant referees (81.8%), compared with their Belgian colleagues (75.1%, $P < 0.001$). In line with the results based on the computer animations, video footage of match-play allowed us to discriminate between standards of expertise. When we examined the results of the FIFA assistant referees in more detail, it was clear that there were large differences between the assistant referees who obtained the highest and lowest scores. In our opinion, a range of 63.2–100% correct is too wide for this elite group. Therefore, it is necessary to fine-tune the decision-making process so that assistant referees make much more uniform and consistent decisions in similar situations. In contrast to the on-the-field test and the offside test based on computer animations, in this test there was no tendency for the assistant referees to raise their flag in cases of doubt.

Based on the results of the present study, it is important to consider the use of an on-the-field offside test as part of the selection procedure of assistant referees. Some aspects are in favour of such a test compared with computer animations. Whereas assistant referees had to assess only the position of moving targets in the computer animations, in the on-the-field offside test an additional difficulty was observed: the assistants had to be positioned exactly on the offside line. Also, the on-the-field test seems to be more representative of match-play, since realistic depth perception and an ability to anticipate the moment at which the ball is passed are part of the on-the-field test.

However, because the computer animations were found to be a good discriminator between the two standards of assistant referees, several reasons emerged to support the off-the-field test as a selection procedure. Regarding the type of errors made by top-class FIFA assistant referees, results clearly showed similarities between the off-the-field and on-the-field offside test. First, there was an overall bias towards flag errors compared with non-flag errors. That bias can be a consequence of the strategy used by assistant referees; top-class assistant referees were more likely to flag in cases of doubt and it is possible that these assistant referees felt psychological pressure to raise their flag. However, in the present study, we found evidence that this strategy of the assistant referees is not sufficient to explain the bias towards flag errors. Forward displacements, associated with the flash-lag effect, were observed in both tests. First, FIFA assistant referees made more flag errors in situations with a dynamic defender than situations with a static defender. Second, in the computer animations, in contrast to the on-the-field test, we were able to manipulate the speed at which the animations were played. We found that FIFA assistant referees made more flag errors when the video footage was played at a speed of 8 frames per second than at 6 frames per second. Based on this result, it can be suggested that at a faster speed, the attacker is perceived to be even further ahead of his position. Third, in both the on-the-field and the off-the-field offside test, the results showed an increase in flag errors when the distance between the attacker in an onside position and the defender decreased.

Not only was the response pattern of the FIFA assistant referees similar in the computer animations and the on-the-field test, a significant correlation was also observed when a comparison was made between the onside situations of the on-the-field test and the computer animations. Consequently, assistant referees who were more accurate in their judgements of onside positions on-the-field were also more likely to make correct onside decisions in the laboratory setting. Similar results were found for the situations with the attacker in an offside position.

The correlation between the response accuracy in the two tests was moderate ($\rho = 0.49$). One explanation is that several assistant referees scored the same percentage for the on-the-field test, so they are given the same rank for that test, while the ranking in the other test differed. A second explanation, which also highlights an important limitation of the on-the-field test, is that it is difficult to control situations in the on-the-field offside test. Some assistant referees might have had to judge more difficult offside situations than other assistant referees, while the difficulty of the computer animations was identical for every assistant referee. Therefore, we cannot be completely confident that a selection of assistant referees based on an on-the-field offside test is completely reliable.

Finally, we evaluated the laboratory and field tests as possible training methods for refining the offside decision-making skills of assistant referees. As stated earlier, structured training for improving knowledge of the laws of the game is important. But even more important during training sessions for assistant referees, the ability to assess the position of moving players needs to be refined. In this regard, we present results that computer animations and video footage of real matches can be complementary training tools, because they clearly discriminate between assistant referees of different standards. We also assume that a transfer to the field will be apparent, because there was a relationship between an on-the-field offside test and an off-the-field test using computer animations. Training offside decision-making skills on-the-field is also a viable option, provided that assistant referees are given immediate feedback using a still frame of the offside situation they just observed and assessed.

Nevertheless, more research with appropriate pre- and post-training test designs is required to test the transfer of different training regimes to performance during match-play by assistant referees in offside decision making. Williams et al. (2002) have devised a training experiment to improve a hockey goal-keeper's anticipation skill at the penalty flick by using video-based perceptual training. In line with this training experiment, a similar setting can be used to test the effectiveness of training offside decisions by the use of computer animations. Along the same lines, the perceptual bias in offside decision making needs to be examined in more detail. Thus, it is important to highlight where the focus of attention of assistant referees is situated while assessing offside. It would be valuable to examine if there is a difference between more and less successful assistant referees with respect to the way they visually scan the offside situation, and where they actually get their information before making an offside call or not.

In summary, the same response pattern was found for FIFA assistant referees when assessing offside situations on the field of play and when presented as computer animations. The results show that more flag errors than non-flag errors were made in the offside decision-making test on-the-field of play and the laboratory offside test based on computer animations. This bias was not present when the assistant referees had only to interpret the offside law. In addition, an expertise effect was found in the laboratory tests, with the FIFA assistant referees performing better. Furthermore, the correlation between the computer animations and the on-the-field test was moderate, but statistically significant. We can conclude that the off-the-field tests of this study clearly discriminated between top-class and national-standard assistant referees. However, it remains premature to advocate use of these tests as the only way to select the most qualified assistant referees from a group of experts. The on-the-field test is much less controllable and is therefore less reliable to select assistant referees than the laboratory tests. Also, training needs to be available for assistant referees to refine their decision-making skills for the two parts of the offside law. Assistant referees are advised to keep their flag down in case of doubt. Further research is necessary to confirm the effectiveness of the on-the-field and laboratory offside tests as a training method that might transfer to match-play.

Acknowledgements

We would like to thank FIFA's F-MARC (Medical Assessment and Research Center) for their financial support in this project. Also, we sincerely thank the members of the FIFA and the Belgian refereeing department for their collaboration in this project and their comments, in particular José-Maria Garçia-Aranda en Fernando Tresaco from the FIFA refereeing department. In addition, the authors are very thankful to the FIFA and Belgian assistant referees for their participation. Finally, we would like to thank Mr. Paul Meugens for designing the research equipment and the electronics.

References

Abernethy, B., & Farrow, D. (2002). Can anticipatory skills be learned through implicit video-based perceptual training? *Journal of Sports Sciences*, *20*, 471–485.

Abernethy, B., & Wood, J. M. (2001). Do generalized visual training programmes for sport really work? An experimental investigation. *Journal of Sports Sciences*, *19*, 203–222.

Baldo, M. V. C., Ranvaud, R. D., & Morya, E. (2002). Flag errors in soccer games: The flash-lag effect brought to real life. *Perception*, *31*, 1205–1210.

Castagna, C., Abt, G., & D'Ottavio, S. (2007). Physiological aspects of soccer refereeing performance and training. *Sports Medicine*, *37*, 625–646.

Ericsson, K. A., Krampe, R. T., & Tesch-Römer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review*, *100*, 363–406.

Ericsson, K. A., & Williams, M. A. (2007). Capturing naturally occurring superior performance in the laboratory: Translational research of expert performance. *Journal of Experimental Psychology: Applied*, *13*, 115–123.

Fédération Internationale de Football Association (2006). *Laws of the game*. Zürich: FIFA.

Finke, R. A., Freyd, J. J., & Shyi, G. C.-W. (1986). Implied velocity and acceleration induce transformations of visual memory. *Journal of Experimental Psychology: General*, *115*, 175–188.

Gilis, B., Helsen, W., Catteeuw, P., & Wagemans, J. (2008). Offside decisions by expert assistant referees in association football: Perception and recall of spatial positions in complex dynamic events. *Journal of Experimental Psychology: Applied*, *14*, 21–35.

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*, 121–138.

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

Helsen, W., Gilis, B., & Weston, M. (2006). Errors in judging "offside" in association football: Test of the optical error versus the perceptual flash-lag hypothesis. *Journal of Sports Sciences*, *24*, 521–528.

Helsen, W. F., & Starkes, J. L. (1999). A multidimensional approach to skilled perception and performance in sport. *Applied Cognitive Psychology*, *13*, 1–27.

Helsen, W. F., Starkes, J. L., & Hodges, N. J. (1998). Team sports and the theory of deliberate practice. *Journal of Sport and Exercise Psychology*, *20*, 12–34.

Hubbard, T. L., & Bharucha, J. J. (1988). Judged displacements in apparent vertical and horizontal motion. *Perception and Psychophysics*, *44*, 211–221.

Krustrup, P., & Bangsbo, J. (2001). Physiological demands of top-class soccer refereeing in relation to physical capacity: Effect of intense intermittent exercise training. *Journal of Sports Sciences*, *22*, 179–189.

Krustrup, P., Mohr, M., & Bangsbo, J. (2002). Activity profile and physiological demands of top-class assistant refereeing in relation to training status. *Journal of Sports Sciences*, *20*, 861–871.

MacMahon, C., Helsen, W., Starkes, J. L., & Weston, M. (2006). Decision-making skills and deliberate practice in elite association football referees. *Journal of Sports Sciences*, *25*, 65–78.

Macmillan, N. A., & Creelman, C. D. (2005). *Detection theory: A user's guide*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Mascarenhas, D. R. D., O'Hare, D., & Plessner, H. (2006). The psychological and performance demands of association football refereeing. *International Journal of Sport Psychology*, *37*, 99–120.

Nijhawan, R. (1994). Motion extrapolation in catching. *Nature*, *370*, 256–257.

Oudejans, R. R. D., Bakker, F. C., Verheijen, R., Gerrits, J. C., Steinbrückner, M., & Beek, P. J. (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.

Oudejans, R. R. D., Verheijen, R., Bakker, F. C., Gerrits, J. C., Steinbrückner, M., & Beek, P. J. (2000). Errors in judging "offside" in football. *Nature*, *404*, 33.

Reilly, T., & Gregson, W. (2006). Special populations: The referee and assistant referee. *Journal of Sports Sciences*, *24*, 795–801.

Sanabria, J., Cenjor, C., Márquez, F., Gutierrez, R., Martinez, D., & Prados Garcia, J. L. (1998). Oculomotor movements and football's Law 11. *Lancet*, *351*, 268.

Starkes, J. L. (1987). Skill in field hockey: The nature of the cognitive advance. *Journal of Sport Psychology*, *9*, 146–160.

Starkes, J. L., Allard, F., Lindley, S., & O'Reilly, K. (1994). Abilities and skill in basketball. *International Journal of Sport Psychology*, *25*, 249–265.

Weston, M., Helsen, W., MacMahon, C., & Kirkendall, D. (2004). The impact of specific high-intensity training sessions on football referees' fitness levels. *American Journal of Sports Medicine*, *32*, 54–61.

Williams, A. M. (2000). Perceptual skill in soccer: Implications for talent identification and development. *Journal of Sports Sciences*, *18*, 737–750.

Williams, A. M., & Davids, K. (1995). Declarative knowledge in sport: A by-product of experience or a characteristic of expertise? *Journal of Sport and Exercise Psychology*, *17*, 259–275.

Williams, A. M., & Grant, A. (1999). Training perceptual skill in sport. *International Journal of Sport Psychology*, *30*, 194–220.

Williams, A. M., & Hodges, N. J. (2005). Practice, instruction and skill acquisition in soccer: Challenging tradition. *Journal of Sports Sciences*, *23*, 637–650.

Williams, A. M., Ward, P., & Chapman, C. (2002). Training perceptual skill in field hockey: Is there transfer from the laboratory to the field? *Research Quarterly for Exercise and Sport*, *74*, 98–103.