Journal of Sports Sciences

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

Offside decision making of assistant referees in the English Premier League: Impact of physical and perceptual-cognitive factors on match performance
Peter Catteeuw \(^a\), Bart Gilis \(^a\), Johan Wagemans \(^b\), Werner Helsen \(^a\)
\(^a\) Department of Biomedical Kinesiology, \(^b\) Department of Psychology, Katholieke Universiteit Leuven, Leuven, Belgium

Online publication date: 23 April 2010

To link to this Article DOI: 10.1080/02640410903518184
URL: http://dx.doi.org/10.1080/02640410903518184

PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: http://www.informaworld.com/terms-and-conditions-of-access.pdf

This article may be used for research, teaching and private study purposes. Any substantial or systematic reproduction, re-distribution, re-selling, loan or sub-licensing, systematic supply or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.
Offside decision making of assistant referees in the English Premier League: Impact of physical and perceptual-cognitive factors on match performance

PETER CATTEEUW1, BART GILIS1, JOHAN WAGEMANS2, & WERNER HELEN1

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

(Accepted 30 November 2009)

Abstract
In the present study, we investigated the accuracy of offside judgements of assistant referees in the English Premier League. The moment in the match, the position and movement speed of the assistant referee, attacker and second-last defender, together with the angle of view for the assistant referee were all considered to underlie incorrect decisions. The error rate was 17.5% (868 of 4960 situations). As the English assistant referees tended not to signal in doubtful situations (c = 0.91), there was an overall bias towards non-flag errors (773 non-flag errors vs. 95 flag errors). The flash-lag hypothesis could explain all flag errors, whereas the optical-error hypothesis could explain a proportion of the non-flag errors (45.4%). Fatigue, movement speed, and angle of view did not have a detrimental effect on offside decision making. In conclusion, there were fewer flag errors than in the 2002 and 2006 FIFA World Cups, whereas the number of non-flag errors rose. The increased awareness of factors involved in offside decision making and the instructions to give the benefit of the doubt to attackers could have contributed to this situation.

Keywords: Association football, assistant referees, offside, decision-making skills

Introduction
The Barclays Premier League in England is probably the biggest club competition in association football (also known as soccer in North America) in terms of financial resources, media coverage, and public attention. In this competition, the outcome of a match can have considerable financial implications for various economic agents involved (e.g. players, staff, clubs, sponsors). Several factors can influence the course and outcome of matches. The refereeing team, which consists of one referee, two assistant referees, and a fourth official, has to make sure that the FIFA (Fédération Internationale de Football Association) Laws of the Game are applied correctly and consistently. The main task of assistants is to make correct offside decisions as prescribed in Law 11. Correct interpretation and application of the offside law can be crucial for the outcome of matches. An incorrect application can result in cancellation of a legal goal or a goal-scoring opportunity, or an illegal goal being allowed to stand.

Law 11 (FIFA, 2007) states that “a player is in an offside position if he is nearer to his opponents’ goal line than both the ball and the second-last opponent.” Any part of the head, body or feet, except the arms, is taken into account when judging if the player is nearer to his opponents’ goal line. A player who is level with the second-last defender, the last two opponents or in his own half of the field of play is not in an offside position. The referee can penalize a player in an offside position only if, at the moment the ball touches or is played by one of the player’s team-mates, he is, in the opinion of the referee, involved in active play by interfering with play, interfering with an opponent or gaining an advantage by being in that position.

Previous research in the highest Dutch football league (Oudejans et al., 2005) and in the 2002 FIFA World Cup in Japan and Korea (Helsen, Gilis, & Weston, 2006) reported error rates for offside decision making of approximately 20% and 26%, respectively. Because of specific technical instructions and training interventions by FIFA to improve the offside decision-making process, the error rate
decreased to 10.0% in the 2006 FIFA World Cup in Germany (Catteeuw et al., submitted). One explanation for this decrease is probably the decision of FIFA to appoint refereeing “teams” for the 2006 World Cup, whereas in 2002 match officials were not accustomed to each other and had different native languages. Second, specific training programmes for assistants were designed for on- and off-field use. Third, selection of assistants was based not only on match reports, as was the case in 2002. Eight weeks before the 2006 World Cup, potential assistants were invited to attend a workshop where a final selection was made, based on the new FIFA fitness tests and on- and off-field offside decision-making tests. Fourth, the referee, his two assistants, and the fourth official used microphone headsets to help facilitate communication and mutual understanding.

To reduce this error rate, it is important to consider all potential factors that could affect the quality of offside decisions. First, fatigue is often cited as a factor leading to incorrect decisions. Krustup and colleagues (Krustup, Mohr, & Bangsbo, 2002) reported deterioration in sprint performance towards the end of matches. This could affect the assistants’ ability to keep up with play because they have to stay as near as possible to the offside line (i.e. the second-last defender) throughout the match. However, Helsen et al. (2006) reported that the number of offside errors did not increase as matches progressed. Second, movement speed could affect the quality of offside decision making. Oudejans et al. (2005) showed that assistants make more mistakes when they are running or sprinting, than when walking or jogging.

Third, assistants might not be able to see the passer and the second-last defender simultaneously (Belda Maruenda, 2004; Sanabria et al., 1998). For instance, when a long pass is made from defence to the attacking third, assistants have to shift their gaze over a large distance, implying that the judgement of an offside occurs a split second after the pass is made. Oudejans et al. (2000), however, found no support for this shift-of-gaze hypothesis. Expert assistants appeared to shift gaze just before the pass, so that they were already looking at the defender at the moment the pass was made.

Fourth, Oudejans et al. (2000, 2005) suggested the optical-error hypothesis to explain incorrect judgements. They differentiated between two types of error: assistants might raise their flag while the attacker is not in an offside position (i.e. a flag error) or, alternatively, they might commit a non-flag error by not raising their flag while the attacker is in an offside position. The type of error can be explained by the inappropriate position of the assistant in relation to the offside line (i.e. trailing or leading it). When assistants are leading the offside line, more flag errors will be made when the attacker is on the opposite side of the second-last defender and more non-flag errors when the attacker is on the near side (Figure 1A). In contrast, when they are trailing the offside line, more non-flag errors are likely when the attacker is on the opposite side of the second-last defender and more flag errors will be found when the attacker is on the near side (Figure 1A).

Finally, Baldo and colleagues (Baldo, Ranvaud, & Morya, 2002) attributed erroneous offside assessments to the flash-lag effect, a visual illusion whereby a moving object is perceived as spatially leading its real position at an instant defined by a time marker (usually a briefly flashed stimulus). In offside situations in football, the attacker receiving the ball can be considered as the moving stimulus, often progressing at high speed. The moment at which the ball is passed can be interpreted as the flash that clearly marks the instant at which the position of the attacker has to be judged by the assistant. As a result, the attacker receiving the ball, often moving forward, is perceived to be ahead of his true position because of the flash-lag effect, leading to more flag errors than non-flag errors (see Figure 1B). In association football, Gilis and colleagues (Gillis, Helsen, Catteeuw, & Wagemans, 2008) showed the existence and consequence of the flash-lag effect using computer animations.

When assistants are not in line with the second-last defender, a combination of the optical-error and flash-lag hypothesis can be used to explain incorrect decisions. However, depending on whether they are behind or ahead of the offside line, an overall bias towards flag errors or non-flag errors will be observed (Figure 1).

The aim of the present study was to investigate offside decision making during the 2007–2008 season of the English Premier League. First, each time the attacker was in an offside position or less than 2 m behind the second-last defender was taken into account. Several variables, including moment during the match, position and speed of the assistant, attacker, and second-last defender, together with the angle of view between passer and second-last defender, were determined for the selected situations.

Second, accuracy was thoroughly investigated and linked with the variables above to evaluate the hypotheses as sources of error. To account for fatigue, both accuracy and movement speed were related to the six 15-min match periods. Because of the introduction of new FIFA fitness tests in 2006, improved fitness can be anticipated for assistant referees. Therefore, in contrast with Krustup et al. (2002), sprinting performances were predicted not to deteriorate and, in line with Helsen et al. (2006), no increase in errors was predicted towards the end of
Figure 1. Distribution of flag and non-flag errors for the optical-error hypothesis (A) and the flash-lag hypothesis (B) in relation to the position of the assistant and attacker. Zones in light grey represent non-flag errors and zones in dark grey represent flag errors.
matches. A negative correlation of movement speed and accuracy, as proposed by Oudejans et al. (2005), was anticipated, meaning more errors when running faster. According to Belda Maruenda (2004) and Sanabria et al. (1998), assistants would probably make a gaze shift, and thus the larger the angle of view the more errors will be made. To account for both the optical-error and flash-lag hypothesis, accuracy was related to the position of the assistant and the two-dimensional position of the attacker. Detailed analysis of all errors was performed to investigate underlying mechanisms in judging offside.

In summary, we examined the factors that can interfere with offside decision making by assistant referees in the English Premier League.

Methods

Participants

Participants were experienced English Premier League top-class assistant referees \((n = 48, \text{ mean age } 41.8 \pm 5.7 \text{ years})\). Ten had been on the FIFA list for 2.6 \pm 3.3 years and had 4.8 \pm 5.1 years’ experience in the English Premier League as an assistant. The other 38 were English elite assistant referees who had 6.6 \pm 5.7 years’ experience as an English Premier League assistant. Written consent was obtained from PGMOL (Professional Game Match Officials Limited), which provides referees and assistants for all professional football matches played in England. ProZone\(^{\text{R}}\) (Leeds, UK), a football match-analysis system, also agreed to cooperate on this project after providing us with their written consent.

Data acquisition

ProZone\(^{\text{R}}\) has an agreement with 14 English Premier League teams to record and analyse their home matches. They provided us with data for 165 of 380 matches played in the 2007–2008 season of the English Premier League. Table I provides an overview of the teams that are affiliated to ProZone\(^{\text{R}}\), and the number of times we obtained data from each team in the English Premier League. ProZone\(^{\text{R}}\) monitored the matches using a computerized, semi-automatic, video-based match-analysis image-recognition system. Eight cameras (four in the corners and four along the side of the stadium) track all players, the ball, the referee, and assistants every one-tenth of a second. Every movement on the pitch was captured because of overlap of the video-footage. After systematic analysis with proprietary software, comprehensive data were provided on the movement of the players, officials, and ball. Recent work (Di Salvo, Collins, McNeill, & Cardinale, 2006; Impellizzeri, Sassi, & Rampinini, 2006) has demonstrated high accuracy and reliability of ProZone\(^{\text{R}}\), suggesting it is valid for practical and research purposes (e.g., Rampinini, Coutts, Castagna, Sassi, & Impellizzeri, 2007; Weston, Castagna, Impellizzeri, Rampinini, & Abt, 2007).

Oudejans et al. (2005) examined those offside situations in which the ball was played forward towards a team-mate who was within a few metres of the offside line. In this study, however, the data set was reduced according to the more precise definition proposed by Catteeuw et al. (submitted), where the attacker is in an offside position or less than 2 m behind the second-last defender (because cases of more than 2 m were regarded as too easy). The objective measures selected for analysis were the precise moment during the match, the position on the field of play, and movement speed of the assistant, attacker, and second-last defender nearest to the offside line. We also obtained the viewing angle between second-last defender and passer from the assistant’s point of view. In addition, we assessed whether the attacker passed the second-last defender on the near side or on the opposite side from the perspective of the assistant.

Data analysis

The Results section is divided into three parts: notational analysis, accuracy, and detailed error analysis.
First, a 2 (situation: onside vs. offside) × 2 (decision: non-flag signal vs. flag signal) contingency table was created. In line with previous studies (Gilis et al., 2008; Gilis, Helsen, Catteeuw, Van Roie, & Wagemans, 2009), assistant referees’ performance was analysed according to signal-detection theory (Macmillan & Creelman, 2005). A sensitivity index (d') was calculated based on the proportion of “hits” (flag signal when the attacker was in an offside position, i.e. correct flag) and “false alarms” (flag signal when the attacker was in an onside position, i.e. flag error). When d' was zero, assistants could not discriminate between onside and offside. In contrast, when d' differed from zero, assistants discriminated between onside and offside. To assess response bias, reported as criterion (Macmillan & Creelman, 2005). In addition, the confidence intervals (95% CI) were calculated when discriminate between onside and offside. In contrast, intensity running (5.5–6.9 m/s) i.e. flag error). When

The angle was divided into six separate categories: 0–15°, 16–30°, 31–45°, 46–60°, 61–75°, and 76–90°. Assuming assistants focused on the second-last defender at the moment of the pass (Oudejans et al., 2000), situations with an angle smaller than 15° are within the foveal field (Palmer, 1999). Situations with an angle larger than 15° are within the peripheral field. Finally, the position of the assistant relative to the second-last defender, and the perpendicular position of the attacker to the touchline and second-last defender were taken into account. In line with Oudejans et al. (2005), assistants were considered to be in line with the offside line if there was a deviation of less than 20 cm. If assistants were not in line, they were either ahead of (i.e. leading) the offside line or behind (i.e. trailing) the offside line. Also to account for potential effects of fatigue, the position of assistants relative to the offside line (Kolmogorov-Smirnov D = 0.079, P < 0.01) was examined in relation to the moment in the match by a Spearman rank correlation and in relation to the six 15-min match periods by a Kruskal-Wallis test. For the position of the attacker, a distinction was made between the perpendicular distance to the touchline and his position relative to the second-last defender. For the perpendicular distance, the field of play was divided into three equal longitudinal strokes (i.e. near, middle or far zone) (see Figure 1). In relation to the second-last defender, the attacker could pass on the near side or on the opposite side from the viewpoint of the assistant.

Chi-squared goodness-of-fit tests were used to examine overall incidence and error rates for these variables.

The detailed error analysis sought possible explanations for flag and non-flag errors. Every error was compared with the optical-error hypothesis (Oudejans et al., 2000, 2005) and with the flash-lag hypothesis (Baldo et al., 2002; Gilis et al., 2008). In Figure 1, an overview is given of the different positions of the assistant and attacker, and the distribution of flag and non-flag errors based on the optical-error and flash-lag hypothesis.

For all analyses, statistical significance was set at P < 0.05.

Results
Notational analysis
First, of 8787 potential offside situations, 3827 were excluded from further analysis because the attacker was more than 2 m behind the offside line. Thus, for 4960 situations from 165 matches (30.1 situations per match, range = 8–60), offside decision accuracy was carefully analysed. In total, 3316 cases of onside (66.9%) and 1644 cases of offside (33.1%) were recorded. Table II shows the characteristics of the onside and offside situations. The assistants flagged
Table II. Characteristics of onside and offside situations.

<table>
<thead>
<tr>
<th></th>
<th>Onside situations</th>
<th>Offside situations</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>s</td>
</tr>
<tr>
<td>Speed of assistant (m·s⁻¹)</td>
<td>2.0</td>
<td>1.5</td>
</tr>
<tr>
<td>Distance of assistant from offside line (m)</td>
<td>0.6</td>
<td>1.5</td>
</tr>
<tr>
<td>Speed of attacker (m·s⁻¹)</td>
<td>3.6</td>
<td>2.0</td>
</tr>
<tr>
<td>Distance of attacker from offside line (m)</td>
<td>-1.1</td>
<td>0.6</td>
</tr>
<tr>
<td>Distance of attacker from assistant (m)</td>
<td>35.8</td>
<td>16.8</td>
</tr>
<tr>
<td>Distance of attacker perpendicular to touchline (m)</td>
<td>35.7</td>
<td>16.9</td>
</tr>
<tr>
<td>Speed of second-last defender (m·s⁻¹)</td>
<td>2.9</td>
<td>1.7</td>
</tr>
<tr>
<td>Distance of second-last defender from assistant (m)</td>
<td>35.3</td>
<td>11.8</td>
</tr>
<tr>
<td>Distance of second-last defender perpendicular to touchline (m)</td>
<td>35.3</td>
<td>11.8</td>
</tr>
<tr>
<td>Angle between passer and second-last defender (°)</td>
<td>28.8</td>
<td>21.8</td>
</tr>
</tbody>
</table>

Table III. Overview of correct and incorrect “flag” and “non-flag” signals.

<table>
<thead>
<tr>
<th></th>
<th>Onside situations</th>
<th>Offside situations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-flag</td>
<td>Correct non-flag</td>
<td>Non-flag error</td>
</tr>
<tr>
<td>Flag</td>
<td>Flag error</td>
<td>Correct flag</td>
</tr>
<tr>
<td></td>
<td>3221 (97.1%)</td>
<td>773 (47.0%)</td>
</tr>
<tr>
<td></td>
<td>95 (2.9%)</td>
<td>871 (53.0%)</td>
</tr>
<tr>
<td>Total</td>
<td>3316</td>
<td>1644</td>
</tr>
</tbody>
</table>

in 966 (19.5%) situations, and they did not flag in 3994 (80.5%) situations.

Accuracy

The assistants correctly assessed 82.5% (4092/4960) of all situations and assessed 17.5% (868/4960) incorrectly (see Table III). For the non-flag errors (n = 773), the median distance of the attacker was 0.7 m (range = 0.1 to 12.7) ahead of the second-last defender. For the flag errors (n = 95), the median distance of the attacker was -0.4 m (range = -2.5 to 0.0) behind the second-last defender.

The assistants discriminated between onside and offside (d = 1.98, 95% CI = 1.87 to 2.08, P < 0.05). The positive c value (c = 0.91) indicated that they did not signal in doubtful and difficult situations.

Accuracy in relation to match periods. The incidence of offside situations was uniformly distributed over the six 15-min periods (χ²(5, N = 4664) = 1.11, P = 0.95). Similarly, correct and incorrect decisions were also uniformly distributed over these periods (χ²(11, N = 4664) = 5.46, P = 0.91) (see Figure 2).

Accuracy in relation to movement speed of assistant, attacker, and second-last defender. First, in most situations the assistants were walking (n = 2658) and then jogging (n = 1,541) (χ²(5, N = 4,960) = 6735.62, P < 0.001). For the remaining situations, they were standing still (n = 234), running (n = 405), running at high intensity (n = 98) or sprinting (n = 24). The error rate, however, was uniformly distributed over the six categories of movement speed (χ²(5, N = 868) = 6.74, P = 0.24). The Kruskal-Wallis test for movement speed of the assistant showed no differences for accuracy (H(3, N = 4,960) = 7.53, P = 0.06). Across all situations, assistants were at walking speed (median = 1.6 m·s⁻¹, range = 0.7–2.8). Using Spearman rank correlation, there was no relationship between moment in the match and speed of the assistant (r = 0.004). The Kruskal-Wallis test for movement speed with match period as the independent variable showed no differences over the six 15-min periods (H(5, N = 4664) = 6.33, P = 0.28).

Second, the attacker was walking (n = 1273), jogging (n = 1736) or running (n = 961) in the majority of situations (χ²(5, N = 4960) = 2438.57, P < 0.001). The errors, however, were uniformly distributed over the six categories of movement speed of the attacker (χ²(5, N = 868) = 9.39, P = 0.09). The Kruskal-Wallis test (H(3, N = 4960) = 20.99, P < 0.0001) revealed that the attacker was running faster when assistants made a flag error (median = 4.8 m·s⁻¹, range = 3.0–5.7) than when they made a correct flag decision (median = 3.3 m·s⁻¹, range = 2.0–4.8), a correct non-flag decision (median = 3.3 m·s⁻¹, range = 1.8–5.0) or a non-flag error (median = 3.2 m·s⁻¹, range = 2.0–4.8).

Third, the defender was walking (n = 1716) or jogging (n = 2036) in the majority of situations (χ²(5, N = 4960) = 4503.71, P < 0.001). The errors, however, were uniformly distributed over the six categories of movement speed of the defender (χ²(5, N = 868) = 10.45, P = 0.06). The defender was running more slowly when assistants made a correct flag decision (median = 2.4 m·s⁻¹, range = 1.5–3.4) than when they made a correct (median = 2.6 m·s⁻¹, range = 1.5–4.0) or incorrect (median = 2.7 m·s⁻¹, range = 1.6–4.1) non-flag decision (H(3, N = 4960) = 14.77, P < 0.01).
incorrect flag decisions, the speed of the defender was similar (median = 2.4 m · s⁻¹, range = 1.4–3.4).

Table IV shows an overview of the distribution of correct and incorrect decisions in relation to the movement speed of the assistant, attacker, and second-last defender.

In addition, movement speed of the assistant and second-last defender was positively correlated ($r_s = 0.51$, $P < 0.05$). Thus, the faster the second-last defender (median = 2.6 m · s⁻¹, range = 1.5–3.9) ran the faster the assistants ran, but on average $T = 2,162,667.00$, $P < 0.01$.

Accuracy in relation to angle of view. In 60.9% of situations, the angle of view of the assistant was between 0° and 15° ($n = 1613$) and 16° and 30° ($n = 1406$). There were fewer situations with an angle of view of 31–45° ($n = 744$), 46–60° ($n = 578$), 61–75° ($n = 426$), and 76–90° ($n = 193$) ($\chi^2(5, N = 4960) = 1916.88$, $P < 0.001$). The errors, however, were uniformly distributed over the six categories of angle of view ($\chi^2(5, N = 868) = 1.96$, $P = 0.80$).

Accuracy in relation to position of assistant and attacker. The assistants were more often positioned ahead of the offside line ($n = 3087$) than behind ($n = 1216$) or on it ($n = 657$) ($\chi^2(2, N = 4960) = 1959.33$, $P < 0.001$).

More errors were made when assistants were positioned ahead of the offside line ($n = 613$) than when they were behind it ($n = 156$) ($\chi^2(2, N = 868) = 27.18$, $P < 0.001$). The proportion of flag errors/non-flag errors was 0.09 when they were ahead of the offside line, and 0.20 when they were behind it. Assistants were more often ahead of the offside line when they made a non-flag error (median = 0.93 m, range = 0.13–1.91) than when they made a correct non-flag decision (median = 0.48 m, range = −0.33 to 1.38), a correct flag decision (median = 0.45 m, range = 0.02–0.85) or a flag error (median = 0.31 m, range = −0.29 to 0.88) ($H(3, N = 4960) = 103.90$, $P < 0.001$).

No correlation was found for the moment of the situation and position of the assistant in relation to the offside line ($r_s = 0.009$). The position of the assistant in relation to the offside line with match period as the independent variable showed no differences ($H(5, N = 4664) = 7.65$, $P = 0.18$).

In relation to the position of the attacker, the offside situations occurred more in the middle part of the field ($n = 2267$) than in the near ($n = 1257$) or far part ($n = 1436$) ($\chi^2(2, N = 4960) = 351.36$, $P < 0.001$). The errors, however, were uniformly distributed ($\chi^2(2, N = 868) = 1.01$, $P = 0.22$). Also, the attacker passed more on the opposite side of the second-last defender ($n = 2623$) than on the near side ($n = 2337$) from the viewpoint of the assistant ($\chi^2(1, N = 4960) = 16.49$, $P < 0.001$). The errors, however, were uniformly distributed ($\chi^2(1, N = 868) = 1.50$, $P = 0.22$).

An overall bias towards non-flag errors was observed. For the combination of the position of assistant and attacker relative to second-last defender, the proportions of flag errors/non-flag errors confirmed the bias (see Table V).

Detailed error analysis

Table VI provides an overview of correct and incorrect “flag” and “non-flag” signals in relation...
to the position of the assistant and attacker. The optical-error hypothesis and the flash-lag hypothesis together explained 446 of 868 (51.4%) errors. While all flag errors could be explained by the flash-lag effect (95/95; 100%) and only a few by the optical-error hypothesis (39/95; 41.1%), more non-flag errors could be accounted for by the optical-error hypothesis (351/773; 45.4%) than by the flash-lag effect (148/773; 19.1%).

**Discussion**

The main objective of this study was to assess the accuracy of offside judgements over one season of club competition. We analysed 165 of 380 matches of the English Premier League (2007–2008 season). All potential offside situations were selected based on flag signals or situations in which the ball was played forwards to a team-mate, who was in an offside position or less than 2 m behind the second-last defender at the precise moment the pass was made. Many offside situations per match were observed (30 on average) compared with just 5 and 7 per match during the 2002 (Helsen et al., 2006) and 2006 World Cup (Catteeuw et al., submitted), respectively. In these studies, however, only flag signals and offside claims by defenders were taken into account. Also, teams in a tournament, like the World Cup, probably play more cautiously to preserve their chances of proceeding to the next round. In a competition like the English Premier League, teams might adopt a more direct, attacking manner of play, which can lead to more potential offside situations.

An overall error rate of 17.5% (868/4960 situations) was found with a clear bias toward non-flag errors (95 flag errors vs. 773 non-flag errors). For the 2002 World Cup, Helsen et al. (2006) reported an error rate of 26.2% and a bias towards flag errors (58 flag errors vs. 9 non-flag errors). For the 2006 World Cup, an error rate of 10.0% was found and assistants made almost as many flag errors ($n = 24$) as non-flag errors ($n = 23$) (Catteeuw et al., submitted). While Oudejans et al. (2005) defined an offside situation as “a situation in which the ball was passed forward towards the right goal and in the direction of a
receiving attacker who was positioned within a few meters of the offside line”, in the present study we used a more precise operational definition. Specifically, situations with the attacker in an offside position or less than 2 m behind the second-last defender at the moment the pass was made were considered potential offside situations and a decision had to be made. In this way, we captured probably every offside decision, whereas the definition of Helsen et al. (2006) and Catteeuw et al. (submitted) underestimated the number of offside decisions and the definition of Oudejans et al. (2005) might have overestimated the number of potential offside situations. The question remains whether an error rate of 17.5% is acceptable and whether there is room for improvement.

Several variables were taken into account in an attempt to explain errors. First, in line with Helsen et al. (2006), the assistants did not make more errors towards the end of matches. The analysis of movement speed according to the moment of the situation and the six 15-min match periods showed no correlation. Movement speed did not deteriorate and had no detrimental effects on the quality of offside decision making towards the end of matches. Also, no correlation was observed between the position of the assistant in relation to the offside line and the moment of the situation. This means that the ability of the assistant to shadow the second-last defender did not deteriorate as the match progressed.

Second, in contrast to the findings of Oudejans et al. (2005), movement speed did not affect accuracy. The assistants made most of their decisions when walking or jogging. Only 10.7% of decisions were made when running faster than 4 m · s⁻¹, but the movement speed of the assistant did not affect the error distribution.

Third, an overall bias towards flag errors could be expected according to the shift-of-gaze theory (Belda Maruenda, 2004; Sanabria et al., 1998), although Oudejans et al. (2000) reported that assistants do not shift gaze at the moment of the pass. They focus on the second-last defender and perceive all necessary information in the peripheral field of view. In this study, the information necessary to judge all offside situations was within the peripheral field of view. Therefore, we assume that assistants focus on the second-last defender and not shift gaze, but further research is necessary to provide support for this contention.

Fourth, according to the optical-error hypothesis (Oudejans et al., 2000), the position of assistant and attacker in relation to second-last defender can have a major impact on accuracy. However, the overall bias towards non-flag errors found here in all cases did not support the optical-error hypothesis. If we consider the perpendicular position of the attacker, the proportions should shift according to Figure 1. But from Table VI, it becomes clear that there were more non-flag errors in every situation.

Table VI. Overview of correct and incorrect “flag” and “non-flag” signals in relation to the position of the assistant, attacker, and second-last defender and potential explanation

<table>
<thead>
<tr>
<th>Position of assistant relative to offside line</th>
<th>Position of attacker relative to touchline</th>
<th>Potential explanation of flag error</th>
<th>Correct flag</th>
<th>Non-flag error</th>
<th>Potential explanation of non-flag error</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Behind</td>
<td>Near</td>
<td>FL/OE</td>
<td>18</td>
<td>25</td>
<td>FL/OE</td>
<td>242</td>
</tr>
<tr>
<td></td>
<td>Opposite</td>
<td>FL</td>
<td>2</td>
<td>2</td>
<td>FL/OE</td>
<td>22</td>
</tr>
<tr>
<td>Middle</td>
<td>Near</td>
<td>FL/OE</td>
<td>33</td>
<td>30</td>
<td>–</td>
<td>277</td>
</tr>
<tr>
<td></td>
<td>Opposite</td>
<td>FL</td>
<td>39</td>
<td>32</td>
<td>OE</td>
<td>331</td>
</tr>
<tr>
<td>Far</td>
<td>Near</td>
<td>FL/OE</td>
<td>8</td>
<td>5</td>
<td>–</td>
<td>42</td>
</tr>
<tr>
<td></td>
<td>Opposite</td>
<td>FL</td>
<td>31</td>
<td>36</td>
<td>OE</td>
<td>302</td>
</tr>
<tr>
<td>On the line</td>
<td>Near</td>
<td>FL</td>
<td>27</td>
<td>14</td>
<td>–</td>
<td>123</td>
</tr>
<tr>
<td></td>
<td>Opposite</td>
<td>FL</td>
<td>7</td>
<td>3</td>
<td>–</td>
<td>23</td>
</tr>
<tr>
<td>Middle</td>
<td>Near</td>
<td>FL</td>
<td>53</td>
<td>20</td>
<td>–</td>
<td>175</td>
</tr>
<tr>
<td></td>
<td>Opposite</td>
<td>FL</td>
<td>43</td>
<td>21</td>
<td>–</td>
<td>167</td>
</tr>
<tr>
<td>Far</td>
<td>Near</td>
<td>FL</td>
<td>4</td>
<td>6</td>
<td>–</td>
<td>17</td>
</tr>
<tr>
<td></td>
<td>Opposite</td>
<td>FL</td>
<td>44</td>
<td>18</td>
<td>–</td>
<td>152</td>
</tr>
<tr>
<td>Ahead</td>
<td>Near</td>
<td>FL</td>
<td>95</td>
<td>146</td>
<td>FL/OE</td>
<td>749</td>
</tr>
<tr>
<td></td>
<td>Opposite</td>
<td>FL/OE</td>
<td>26</td>
<td>16</td>
<td>–</td>
<td>98</td>
</tr>
<tr>
<td>Middle</td>
<td>Near</td>
<td>FL</td>
<td>119</td>
<td>120</td>
<td>OE</td>
<td>611</td>
</tr>
<tr>
<td></td>
<td>Opposite</td>
<td>FL/OE</td>
<td>168</td>
<td>111</td>
<td>–</td>
<td>706</td>
</tr>
<tr>
<td>Far</td>
<td>Near</td>
<td>FL</td>
<td>25</td>
<td>15</td>
<td>OE</td>
<td>101</td>
</tr>
<tr>
<td></td>
<td>Opposite</td>
<td>FL/OE</td>
<td>129</td>
<td>153</td>
<td>–</td>
<td>822</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td>871</td>
<td>773</td>
<td></td>
<td>4960</td>
</tr>
</tbody>
</table>

#OE = optical error; FL = flash-lag.
Fifth, according to the flash-lag hypothesis (Baldo et al., 2002; Gilis et al., 2008; Helsen et al., 2006), the attacker receiving the ball, who is often moving forwards, is perceived to be ahead of his true position, leading to more flag errors than non-flag errors. However, an overall bias towards non-flag errors was observed. It is likely that the assistants were following Law 11 more stringently than before, only flagging very clear and unambiguous offside situations.

Keeping in mind the position of assistant and attacker, only 41.1% of flag errors could be explained by the optical-error hypothesis. Interestingly, all flag errors could be explained by the flash-lag effect. Furthermore, the optical-error hypothesis could account for 45.4% of non-flag errors, while the flash-lag hypothesis could only explain 19.1%. The signal-detection analysis might explain the remaining 54.6% (422/773) non-flag errors. The assistants tended not to signal in doubtful and difficult situations (c = 0.91).

To account for the clear bias towards non-flag errors, the training and guidance of the English assistant referees needs to be taken into account. First, after the 2002 World Cup, FIFA decided to take some initiatives to improve the performances of assistants and to decrease the number of errors in judging offside. FIFA worked with fixed trios for the 2006 World Cup, they held specific on-field and off-field selection tests for assistants, and they also introduced microphone headsets to improve the communication between the referee and two assistants. Specific training programmes were created to better deal with the flash-lag phenomenon on the field of play, including immediate video-feedback after assessing real live offside situations, which probably led them to flag only if there was a clear offside situation. Second, the specific guidance of assistants in England, not to signal in case of doubt, could also have had a major impact. At the start of the season, the assistant-referee manager in the English Premier League stated: “There is an emphasis these days on delaying the flag and giving the benefit of the doubt to the attacking team” (Rejer, 2007). As a consequence, more attacks can continue and probably more goals will be scored, which can make the game more attractive to watch. In this study, the assistants did not signal in case of doubt. To improve application of the offside law across matches and competitions, perhaps PGMOL should consider being more explicit with respect to the distance between attacker and second-last defender that results, or not, in a doubtful situation.

To improve accuracy of offside judgements, several training approaches should be considered. First, on-field offside decision-making training can be beneficial. Thus, players should be instructed to act as defenders and attackers when simulating offside situations on the field. In addition, the use of immediate video-feedback is essential to allow assistants to compare their perception with the real positions of the players in the still frame at the precise moment the ball was played. Second, off-field training tools have been shown to discriminate between higher- and lower-standard assistants (Gillis et al., 2009). Consequently, off-field offside decision-making tests and exercises with computer animations and video simulations clearly can aid in on-field offside test performance. In this off-field training, again, instant feedback is essential to let assistants compare their perception with the exact position of players in the still frame. Nevertheless, transfer to the match and training benefits need to be investigated in future research. Earlier studies addressed these issues in different fields. Williams and colleagues (Williams, Ward, & Chapman, 2002) showed the effectiveness of a video-based perceptual training programme in field hockey. Devos et al. (2009) found training in a driving simulator to improve driving skills after stroke.

Further research is also needed to gain more insight in the direction of attention of assistants when judging offside. Previously, it has been shown that the visual attention strategies of experts differ from those of beginners because they have the capability to discriminate between relevant and irrelevant information in task-specific skills (Abernethy & Russell, 1987; Williams, Davids, & Williams, 1999). In this respect, the examination of eye movements and visual search strategies can be a central topic in future offside decision-making research.

In conclusion, the type of errors made has tipped towards non-flag errors rather than flag errors as in previous studies (Catteeuw et al., submitted; Helsen et al., 2006; Oudejans et al., 2005). Fatigue, movement speed, and angle of view did not seem to be viable explanations for incorrect decisions. The optical-error and flash-lag hypothesis are able to account for all flag errors and some of the non-flag errors. The guidance and instructions by international federations (FIFA and UEFA) and national associations, here PGMOL, probably had an important impact on erroneous offside judgements. The instruction not to signal in case of doubt and increased awareness of assistants regarding the underlying mechanisms in offside decision making might have resulted in overcompensation. Thus, flag errors fell to an acceptable rate, whereas the number of non-flag errors rose.

References


