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

Helsen, Gilis and Weston (2006) err in testing the optical error hypothesis

Raôul R. D. Oudejans[a]; Frank C. Bakker[a]; Peter J. Beek[a]
[a] Institute for Fundamental and Clinical Human Movement Sciences, Vrije Universiteit, Amsterdam, The Netherlands

First published on: 08 March 2007

To cite this Article Oudejans, Raôul R. D. , Bakker, Frank C. and Beek, Peter J.(2007) 'Helsen, Gilis and Weston (2006) err in testing the optical error hypothesis', Journal of Sports Sciences, 25: 9, 987 — 990, First published on: 08 March 2007 (iFirst)
To link to this Article: DOI: 10.1080/02640410600778610
URL: http://dx.doi.org/10.1080/02640410600778610

Journal of Sports Sciences, July 2007; 25(9): 987 – 990

Routledge
Taylor & Francis Group

Helsen, Gilis and Weston (2006) err in testing the optical error hypothesis

RAÔUL R. D. OUDEJANS, FRANK C. BAKKER, & PETER J. BEEK

Institute for Fundamental and Clinical Human Movement Sciences, Vrije Universiteit, Amsterdam, The Netherlands

(Accepted 3 April 2006)

Abstract
In this commentary, we react to the recent study by Helsen, Gilis and Weston (2006) on judging offside in football. Helsen *et al.* claim that their data falsify the optical error hypothesis presented by Oudejans *et al.* (2000). However, as we will elucidate here, they misinterpret this hypothesis and present a data set that is seriously flawed, and hence not suited to test it. Therefore, their conclusions regarding the optical error hypothesis are in error.

Keywords: *Assistant referee, football, judging offside, refereeing*

Introduction

In a recent study, Helsen, Gilis and Weston (2006) address several relevant aspects of offside judgements in association football based on an analysis of all 64 matches at the 2002 World Cup. Such studies are important for gaining further insight into refereeing behaviour and to identify explanations for incorrect offside decisions. Several of Helsen and colleagues' findings add to the existing knowledge base on offside decisions, including the fact that the distance between the receiving attacker and the offside line is greater for correct than for incorrect decisions. Apparently, when the attacker is further away from the offside line, it is easier to judge whether he is positioned on- or offside (see also Oudejans *et al.*, 2005).

However, the study by Helsen *et al.* (2006) also suffers from major shortcomings that prompted us to write this commentary. A central aim of the study was to test two explanations for errors in judging offside that have been proposed in the literature, namely the optical error hypothesis (Oudejans *et al.*, 2000) and the perceptual flash-lag hypothesis (Baldo, Ranvaud, & Moyra, 2002). Helsen *et al.* conclude that their results refute the optical error hypothesis by Oudejans *et al.* (2000) on four counts (p. 527). Unfortunately, as we will elucidate here, none of these counts is properly substantiated due to incorrect interpretations of the optical error hypothesis and the use of a data set that is not suited to test it. Before critically discussing the study of Helsen *et al.*, we will briefly explain the optical error hypothesis.

The optical error hypothesis by Oudejans *et al.* (2000, 2005)

According to the optical error hypothesis, (expert) assistant referees use a variable for judging offside that does not always specify actual relative player positions, namely the optical angle between the second last defender and the receiving attacker. This angle only correctly specifies who is closer to the defender's goal line (attacker or defender) when the assistant referee is positioned on the offside line. When the assistant referee is so positioned, a negative angle between defender and attacker (implying that the attacker is positioned further towards the halfway line than the defender from the assistant's perspective) or a zero angle specifies that the attacker has not gone past the defender, whereas a positive angle (implying that the attacker is positioned closer to the goal line than the defender) specifies that the attacker has gone past the defender and is thus offside. When the assistant referee is not positioned on the offside line (see Figure 1), this angle no longer correctly specifies whether or not the attacker is in an offside position, leading to a

Correspondence: R. R. D. Oudejans, Faculty of Human Movement Sciences, Vrije Universiteit, Van der Boechorststraat 9, 1081 BT Amsterdam, The Netherlands. E-mail: r.oudejans@fbw.vu.nl

ISSN 0264-0414 print/ISSN 1466-447X online © 2007 Taylor & Francis
DOI: 10.1080/02640410600778610

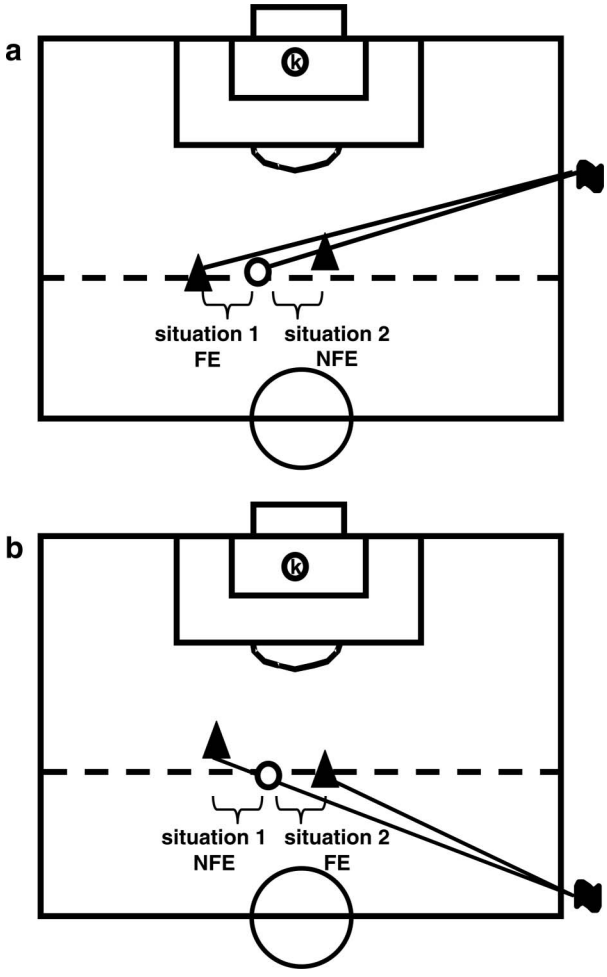

Figure 1. Top view of the optical angle between the attacker (▲), the assistant referee (black flag), and the defender (○) when the assistant referee is leading (a) or trailing (b) the offside line. FE (flag error) and NFE (non-flag error) represent the types of error that are expected to prevail in each situation. In Situation 1 there is a defender between the attacker and the assistant, whereas in Situation 2 there is not. The open circle with the "k" represents the goalkeeper. [Reprinted with permission from Oudejans *et al.* (2005).]

predictable pattern of errors that depends on the position of the assistant referee (behind or ahead of the offside line) and the relative positions of the relevant players (i.e. defender and attacker). Indeed, Oudejans *et al.* (2000, 2005) showed that the assistant referees were frequently positioned away from the offside line when they were judging offside, and occupied a point of observation from which errors were optically probable. Furthermore, they showed that the errors made in judging offside were consistent with the use of the information source just described. Oudejans *et al.* (2000, 2005) therefore concluded that the observation point of assistant referees relative to the offside line is an important determinant of incorrect decisions in judging offside.

Failures in falsifying the optical error hypothesis

Asymmetrical error distributions

On page 527 of their article, Helsen *et al.* (2006) summarize the four grounds on which they reject the optical error hypothesis. First, they argue that their finding of more flag errors than non-flag errors is inconsistent with this hypothesis. Flag errors occur when the assistant referee flags while the receiving attacker is not offside and non-flag errors are made when the assistant referee does not flag while the receiving attacker is offside. Helsen and colleagues' idea that more flag than non-flag errors refutes the optical error hypothesis is related to their claim that Oudejans *et al.* (2000) "predicted a symmetric phenomenon in which flag errors and non-flag errors should occur with equal probability if the positions of the attacking player are equally spread on the opposite and the near side of the second last defender" (Helsen *et al.*, 2006, p. 523). This claim is incorrect. As explained in the preceding, the critical prediction emanating from the explanation by Oudejans *et al.* (2000) is that the type of error (i.e. flag error or non-flag error) depends on the positioning of relevant players (i.e. attacker and defender) as well as that of the assistant referee. Contrary to Helsen and colleagues' claim, Oudejans *et al.* did not predict that the distribution of flag and non-flag errors would be symmetrical because their hypothesis does not pertain to the relative positioning of relevant players and the assistant referee, only to the optical consequences of that positioning. One reason for an asymmetry in flag and non-flag errors that is consistent with the optical error hypothesis is that the farther the relevant players are positioned away from the assistant referee, the more difficult it will become to judge offside, as it is more difficult to detect differences in depth between players, consequently leading to more errors when the relevant players are far from the assistant referee compared with when they are near him (see Oudejans *et al.*, 2005). The findings of Oudejans *et al.* (2000; 266 errors "far" vs. 156 errors "near") are consistent with this idea.

Furthermore, as argued by Oudejans *et al.* (2005), it is more likely in general that one or more defenders are positioned between the receiving attacker and the assistant referee (Situation 1 in Figure 1) rather than none (Situation 2). Only when the attacker is near the assistant referee is the latter situation more likely. Therefore, we never predicted a symmetrical pattern of errors. Instead, we predicted that assistant referees would make *relatively* more flag errors than non-flag errors when they are leading the offside line, and that they would make *relatively* more non-flag errors than flag errors when they are trailing the offside line. The pattern of errors that we found

confirmed those predictions (Oudejans *et al.*, 2000, 2005). In summary, the symmetry test performed by Helsen *et al.* (2006) is not an appropriate test of the optical error hypothesis forwarded by Oudejans *et al.* (2000, 2005).

Even if the (a)symmetry in the distribution of flag and non-flag errors could provide an adequate test of the optical error hypothesis, the data set presented by Helsen *et al.* (2006) is seriously flawed and not suited to test it. Although Helsen *et al.* analysed all matches played at the 2002 World Cup, an important limitation is that all observations were obtained from television images, which reduced the data set from an initial sample of 337 to 256 situations for most of the analyses and even to 61 situations for the ''in-depth'' analysis of positioning of relevant players and assistant referees. The consequences of these reductions for the results and conclusions are unclear, especially in terms of the distribution of flag and non-flag errors.

Most importantly, Helsen *et al.* (2006) drew selectively on the available television images resulting in a data set that is partly subjective (as we will demonstrate shortly) and markedly skewed, rendering conclusions about symmetry of error distributions impossible. Specifically, they made a distinction between two types of decisions by the assistant referees – namely, when they flagged (and gave offside) and when defenders claimed offside but the assistant referee did not flag (offside was not given). Scoring the first type of decision clearly captures all flag situations. But scoring only the non-flag situations that were claimed by defenders yields a small subjective subset of all non-flag situations. That the subset in question represents a subjective selection is underscored by Helsen and colleagues' own finding that in 73.5% of the selected non-flag situations, the claims by the defensive players were incorrect (see their Table I, p. 525). That it only involves a small subset of all potential non-flag situations becomes clear in light of the findings of Oudejans *et al.* (2005) on potential offside situations in competitive matches.

Oudejans *et al.* (2005) defined a potential offside situation as that in which the ball was passed forward towards the goal and in the direction of a receiving attacker who was positioned within a few metres of the offside line. As a consequence, assistant referees had to make a decision about offside in such situations. Oudejans *et al.* scored all potential offsides (at the right side of the field) of four real matches using video recordings made for the specific purposes of that study. In total, 215 situations were selected, of which 194 (90%) were non-flag situations and 21 (10%) were flag situations. Helsen *et al.* (2006) selected 222 flag situations. Assuming that there were no large deviations in the proportion of flag and non-flag situations compared with the proportion found by Oudejans *et al.*, this would amount to a complete population of about 2200 potential offside situations and about 2000 non-flag situations. This implies that by only scoring the offsides claimed by defenders but not given by the assistant referees ($n = 34$), Helsen *et al.* effectively confined their analyses to less than 2% (34 of 2000) of all non-flag situations. As a result, they must have missed non-flag errors in their selection, which is confirmed by the fact that in none of the five non-flag errors reported by Oudejans *et al.* (right half of the pitch), was offside claimed by one or more defenders. Thus, given the limited sample of non-flag situations in their data set, it is hardly surprising that Helsen *et al.* found asymmetries in the error distribution.

Errors on the near side of the second last defender

The second count concerns the finding by Helsen *et al.* (2006) that there were more flag errors than non-flag errors on the near side of the second last defender. Assuming that ''the near side'' refers to the position of the receiving attacker involved in the offside situation relative to the second last defender, it is evident from point 1 (see above) that the underrepresentation of non-flag situations in their sample readily explains why more flag errors than non-flag errors were found on the near side of the second last defender. In addition, according to the optical error hypothesis, whether more flag or non-flag errors are expected when the receiving attacker is further away from the assistant referee than the defender (Situation 1 in Figure 1), compared with when the attacker is positioned between the defender and assistant referee (Situation 2), depends on whether the assistant is trailing or leading the offside line (compare Figures 1a and 1b). Oudejans *et al.* (2005) found that in the four competitive matches analysed, assistant referees trailed the offside line in 54% of cases, led the offside line in 33%, and were in line in 13%. Thus, whether assistant referees are ahead of or behind the offside line is crucial for the type of error to be expected. Unfortunately, Helsen *et al.* arrived at their conclusions without analysing the decisions made by the assistant referee as a function of their position relative to that of relevant players. As a consequence, their analyses are inconsequential for the optical error hypothesis.

Unexpected errors on either side of the second last defender

Third, Helsen *et al.* (2006) found that ''on the opposite side of the second last defender there were also non-flag errors, while flag errors also occurred

on the near side of the second last defender'' (p. 527). Again, we presume that by ''the opposite and near side'' the authors mean the position of the receiving attacker relative to the defender and assistant referee. Why this finding would undermine the optical error hypothesis, as the authors suggest, is unclear, given that such errors are specifically predicted by the hypothesis when the assistant is trailing the offside line, as can be seen in Figure 1b. Moreover, Oudejans *et al.* (2000) had previously reported such errors together with their original presentation of the hypothesis. Once more, whether flag or non-flag errors are expected not only depends on the positioning of the relevant players relative to one another and the offside line, but also on the positioning of the assistant referee relative to the offside line (Figures 1a and 1b).

Distance of assistant referees to the offside line

Finally, according to Helsen *et al.* (2006) it is also problematic for the optical error hypothesis that, on average, the position of the assistant referee relative to the offside line was similar for correct and incorrect decisions (0.81 and 0.77 m ahead of the offside line, respectively). Again, we fail to see why this would be the case. It does not follow logically from the optical error hypothesis that the further away assistant referees are positioned from the offside line, the more errors they will make. Importantly, the hypothesis is likely to apply when assistants are unaware of their poor positioning. In this context, it is noteworthy that the assistant referees examined by Oudejans *et al.* (2000) indicated, in informal interviews afterwards, that they were indeed unaware of their poor positioning, believing that they were in line with the second last defender. Being offline by a large distance may inform the assistant that he or she is not on the offside line, leading to a different perceptual basis for making the decision. As we have already emphasized, whether being positioned ahead of or behind the offside line leads to errors also depends on the positions of the relevant players. Therefore, the optical error hypothesis can only be tested by considering the combination of decision and positioning of relevant players as well as the positioning of the assistant referee in a large, unbiased sample of offside situations. The analyses performed by Helsen *et al.* do not meet these criteria, and thus lack the power to falsify or verify the optical error hypothesis. Note that by implication the study also lacks the power to test the perceptual flash-lag hypothesis, especially as Helsen *et al.* did not report the number of incidents where the attacker was running towards the goal, towards the midline, or standing still when offside judgements were made. This information is necessary to draw conclusions about the possible contribution of flash-lag effects to errors in judging offside.

In summary, Helsen *et al.* (2006) erred in testing the optical error hypothesis as forwarded by Oudejans *et al.* (2000) because their interpretation of the hypothesis was inadequate and because they did not examine a representative sample of the entire population of all potential offside situations in competitive matches. Instead, they tested a skewed and very limited sample of offside situations. The fact that the non-flag situations were selected on the basis of subjective and selective claims of defenders undermines any conclusion about the specific hypotheses tested, not least because it resulted in an investigation of less than 2% of relevant non-flag situations.

Acknowledgements

We thank two anonymous reviewers for helpful comments on an earlier draft of this paper.

References

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

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.

Oudejans, R. R. D., Bakker, F. C., Verheijen, R., Steinbrückner, M., Gerrits, J. C., & 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* (6773), 33.