INFLUENCE OF REACTION TIME AND MOVEMENT IN THE PERFORMANCE OF FOOTBALL REFEREES

INFLUENCIA DEL TIEMPO DE REACCIÓN Y DE MOVIMIENTO EN EL RENDIMIENTO DE ÁRBITROS DE FÚTBOL

RICARDO DE-LA-VEGA¹, DANIEL CARRASQUILLA¹, ENRIQUE ORTIZ², ROBERTO RUIZ¹ Y MANUEL ARMENTEROS³

Abstract

The primary objective of this research is to study the relationship between reaction time (RT) and movement time (MT) in a sample of 121 football referees. Assessment has been performed using an objective computerised test (RT, Vienna Test System⁶), with regard to the average score obtained for refereeing performance over the season (RAR in Spanish). To perform the analysis, the following variables have been considered: a) Level refereed (2 levels); b) Age (3 levels); and c) Experience (4 levels). The hypothesis states that: (i) The higher the level refereeing at, better results are expected in RT and MT; (ii) The older the referee, the greater the expected value in RT and MT; (iii) The more experienced the referee, the lower the score expected in RT and MT; and (iv) It is expected that there will be a negative correlation

Acknowledgments: We are grateful to the company Schuhfried© which has offered its teams for research to be carried out with football referees by this research team. In addition, we would like to express our gratitude to the Technical Committee of Referees of the Royal Football Federation of Madrid and the Royal Football Federation of Spain.

Corresponding author: Ricardo de la Vega Marcos. Departamento de Educación Física, Deporte y Motricidad Humana. Universidad Autónoma de Madrid. Email: ricardo.delavega@uam.es

ORCID: Ricardo de la Vega (http://orcid.org/0000-0002-7395-3297) y Manuel Armenteros (http://orcid.org/0000-0002-0120-908X).

¹Universidad Autónoma de Madrid, España.
²Real Federación de Fútbol de Madrid, España.
³Universidad Carlos III, España.

Recibido: 30 de enero de 2017.
Aceptado: 12 de abril de 2017.
between the RAR and the RT and MT scores. The results show that: (i) RT and MT do not differentiate between semi-professionals and amateur referees; (ii) Only the older group has differences in line with the proposed hypothesis; (iii) There are no differences in RT and MT regarding refereeing experience; (iv) The RT scores, taking the optimal transformation parameters, show significant correlations with the RAR ($r = -.271$, $p = .020$); but not regarding the MT and RAR. Future research should take a deeper look into the results obtained, including professional referees in the comparisons, as well as investigating the effects of other psychological variables in refereeing performance.

**Keywords:** reaction time; movement time; Vienna Test System; refereeing; football.

### Resumen

El objetivo prioritario de esta investigación se centra en estudiar, en una muestra de 121 árbitros de fútbol, la relación que existe entre su tiempo de reacción (TR) y de movimiento (TM) -evaluado mediante la realización de una prueba objetiva computerizada (RT, Vienna Test System®)-, respecto a la puntuación promedio obtenida durante la temporada sobre su rendimiento arbitral real (RAR). Para realizar los análisis se han considerado las siguientes variables: (a) **Categoría arbitral** (2 niveles); (b) **Edad** (3 niveles); y (c) **Experiencia** (4 niveles). Las hipótesis establecen que: (i) A mayor categoría arbitral, se esperan mejores resultados en TR y TM; (ii) A mayor edad, se esperan obtener mayores valores en TR y TM; (iii) A mayor experiencia, se esperan menores puntuaciones en TR y TM; y (iv) Se espera que exista correlación negativa entre RAR, con sus puntuaciones en RT y en TM. Los resultados muestran que: (i) TR y el TM no permiten discriminar entre los árbitros semi-profesionales y los amateurs; (ii) Sólo el grupo de más edad tiene diferencias en la línea de la hipótesis plantead; (iii) No se producen diferencias en RT y TM en función de las experiencias arbitrales; y (iv) Las puntuaciones en RT, tomando el parámetro óptimo de transformación, muestran correlaciones significativas con RAR ($r = -.271$, $p = .020$); no así respecto a TM y RAR. Futuras investigaciones deben profundizar en los resultados obtenidos, añadiendo en las comparaciones a los árbitros profesionales, además de explorar el efecto de otras variables psicológicas en el rendimiento arbitral.

**Palabras clave:** tiempo de reacción; tiempo de movimiento; Vienna Test System; arbitraje; fútbol.

### Introduction

Among the work carried out for quite some time by the Technical Committee of Referees of the Royal Football Federation of Madrid, the promotion and development of refereeing talent in several disciplines has an important role. This follows recommendations established by the *Fédération Internationale de Football Association* (FIFA). Within these areas, the drive to include analysis and development of different psychological variables which may influence refereeing performance has formed a constant part of the work done within this Federation. In this context, in which approximately 1,200 referees control games every weekend in the Community of Madrid, it was proposed to respond to the assessment requirements of different refereeing skills. After several meetings, the following skills were considered a priority: reaction time, stress response, anticipation of space and time and concentration (Spitz, Put, Wagemans, Williams, & Helsen, 2016).

Just as the physical condition of athletes has significantly increased in modern football, demanding greater strength, resistance and speed on the part of the players, the growing requirements and demands of football have affected the preparation requirements of referees (Blumenstein & Orbach, 2014). Their work demands specific attributes which are different from other athletes, being much more abundant from a psychological point of view. Referees work under stress and pressure both on and off the pitch. They are responsible for the implementation of the rules and that players follow these meetings, the following skills were considered a priority: reaction time, stress response, anticipation of space and time and concentration (Spitz, Put, Wagemans, Williams, & Helsen, 2016). It is necessary to understand how complex refereeing is, as they must evaluate and
make quick decisions, maintaining order and the flow of the match, while all the time being challenged and questioned by players, coaches, spectators, the press, and so on. Not forgetting that this is all done in a short period of time (Guillén & Feltz, 2011). The results of the study carried out by Guillén, Moran, and Castro (1997) are interesting in this respect. They discovered that referees recognised basic deficiencies in their training, especially on a psychological level. This was also observed in research performed by González and Dosil (2005), in which the scarce number of hours referees received training on psychological factors is confirmed, despite it being considered essential for their work. In line with the approaches identified by Caracuel, Andreu, and Pérez (1995), referees must perform three basic tasks to reach a correct decision: notice the situation which requires judgement; compare the situation with the rules of the game; and take appropriate action. The speed of the decision-making process is essential and, therefore, also the reaction time from the incident occurring to them making the decision. In addition, this decision is expected to be correct and appropriate.

Considering that the decision-making process for referees depends on their perception of the relevant physical stimuli and the selection and organisation of these, it can depend to a large extent on the specific nature of the task they are faced with (Robles, 2014; Vargas, 1994). In the case of refereeing in football, the quality of the decision-making is affected by the understanding of adequate discriminatory stimuli which occur in a particular situation, with them also having to assess the different calls they can make in a particular situation which demands they also do it in the shortest time possible. This justifies the importance of reaction time analysis as a relevant variable of study regarding referees. Given the speed in which action occurs in a match due to the interaction between teammates and opponents on the pitch, the visual ability and the reaction time of the referee take on special importance (Guillén et al., 2003; Rudzitis, Kalejs, & Licis, 2014).

There are many factors which influence the reaction time in a sporting context. They can be classified as those related to factors which depend on the subject and independent factors (those connected to stimulus). There are also those which correspond to the different perception and processing system, and the response system (Pérez, Soto, & Rojo, 2011). Within the factors dependent on the subject, there is the initial physical condition, the performing or not of pre-match warm-up, the modulating role of tiredness, motivation, dehydration, the body part which performs the response, as well as other characteristics particular to the subject such as age, gender, substances taken (stimulants for the central nervous system such as caffeine), type of sport played and sporting level (Henry & Rogers, 1960; Robles, 2014; Sage, 1977). On the other hand, there are the factors related to the stimulus. This includes its physical characteristics, intensity, means of transmission, period leading up to it, complexity of movement and the influence of colour in the stimulus (Famose, 1992; Schimdt, 1975).

Focusing our interest on referees, the studies of Put, Baldo, Cravo, Wagemans, and Helsen (2013), have shown that factors such as brightness or the peculiarity of the stimulus can cause a ‘flash lag effect’ on the referee, which is an optical illusion in which offside is incorrectly given. For FIFA and UEFA, as leading bodies of football at an international and European level, previous experience and recorded learning are prerequisites to referees performing at the highest level (Castagna, Abt, D’Ottavio, & Weston, 2005). According to the study by Neville, Balmer, and Williams (2002), referees with more experience are more likely to have greater control over their emotions and their particular tasks, improving their decision-making skills under stress. Along those lines, Caracuel et al. (1994), perform the psychological analysis of refereeing using the Interbehavioral Model in which certain personal attributes are highlighted such as the fact that experience could influence on perception and as a result, the reaction time of the referee.

With regard to the maturity in terms of ideal age for the highest performance of refereeing, Stamford (1988) states that it is a multifactorial process that can have a profound effect on human performance. Castagna et al. (2005) state that referees normally reach their highest level of performance when they are close to 40. In their relationship with RT, the drop in performance is connected to the ageing process which brings about the slowing down of the transmission of the neural electrical
Interesting study on 18 referees at an international level (Rudzitis et al., 2014), and Ghasemi, Momeni, Rezaee (2014) for this group in a previous research study carried out by Taheri (2011), compare the sample of 32 referees, Ghasemi, Momeni, Jafarzadehpur, Rezaee and Gholami (2009), on their part, carried out an interesting study on 18 referees at an international level in different team sports, assessing the different psychological skills, among which it is worth highlighting in particular the RT assessed using the Vienna Test System. The results show two differential factors in top-level referees: the speed of the decision-making and its quality (as a percentage of accurate decisions compared to lower-level referees). Within the conclusions, they emphasised the importance of new research on these aspects in specific sports and the need for training in these variables to improve the quality of the training for referees.

Giving the interest in studying the role RT and MT have on football referees, the aim of this study is focused, on one hand, on analysing if the refereeing level, age, experience and physical condition have a direct influence on the RT and MT of football referees. On the other hand, it is interested in finding out if there is a relationship between the scores obtained in RT and MT with the refereeing performance evaluated by the Technical Committee of Referees (CTA in Spanish) via observers who are responsible for this. The hypotheses suggested are: (i) The higher the level refereeing at, the better the results expected in RT and MT; (ii) The older referee, the greater the value is expected in RT and MT; (iii) The more experienced, the lower the score expected in RT and MT; and (iv) It is expected that there is a negative correlation between the RAR and the RT and MT scores.

Method

Participants

All the participants in the sample who have taken part in this research project are holders of the qualification to be a football referee awarded by the Royal Spanish Football Federation through the Royal Football Federation of Madrid. They are all currently performing this work at their corresponding level, meaning that they are all active. This work has had the invaluable collaboration from the Technical Committee of Referees (CTA) of the Royal Football Federation Madrid, with a schedule es-
tablished with the most optimal times the tests to be performed.

The study sample is made up of 120 referees \((M_{\text{age}} = 26.9; SD = 7.35)\), chosen incidentally and confidentially, with them giving their prior consent and complying with the ethical requirements of the Declaration of Helsinki.

**Instruments**

For the assessment of the RT and MT, the RT/S3 test has been used for the psychological diagnostic system, the Vienna Test System \(\text{VTS}\) which belongs to the company Schuhfried (http://www.schuhfried.com). The accuracy of the test is of \(\alpha_{\text{RT}} = .937; \alpha_{\text{MT}} = .979\). The test gives excellent values for convergent validity (Karmer & Neuwirth, 2000).

Through the realisation of the RT test, the following scores were obtained: (a) average RT (ms); (b) Average MT (ms); (c) Box Cox distribution of the RT, which allows distribution biases of the errors made to be corrected; (d) Box Cox distribution of the MT, which allows distribution biases of the errors made to be corrected.

For the assessment of the refereeing performance (RAR), the CTA has provided the average scores obtained by each of the referees involved in the study, while respecting the principle of confidentiality of the study. This is obtained with a minimum of three matches monitored by specially dedicated experts, using a scoring scale from 0-10, in which 10 means that the performance has been excellent, while 0 represents a very bad performance. This information is supplied on a weekly basis to the CTA and the results obtained during the study have been used at the end of the normal 2015/2016 season.

**Procedure**

All the assessments were made in the CTA facilities of the Royal Football Federation of Madrid and in optimal environmental conditions for the adequate performance of the tests. Each of the participating referees had to sign an obligatory informed consent which stated the schedule given by the CTA on the Internet, so that the exact times for each of the applications were established.

The Reaction Time (RT) test of the Vienna Test System \(\text{VTS}\) consists of moving a bone of the finger located on a metallic plate towards a target key depending on the appearance of a specific stimulus (sharp tone at 2000 Hz and yellow colour simultaneously). The test begins with some brief instructions which describe the task. Subsequently, there is a short practice phase in which at least 9 stimuli appear, which is used to check that the subject has understood the task. Once this phase has been completed, the experimental phase begins. This includes a total of 48 stimuli, of which 16 are correct answers. The time required for administration of the test (including instructions) is approximately 9 minutes.

**Independent variables**

For this study, variables of age, experience and level being refereed at are all considered with the aim of being able to establish if there are significant differences in interaction with the dependent variables: reaction time and movement time.

**Age.** \((M_{\text{age}} = 26.9; SD = 7.35)\). For a better analysis of the data, the sample is divided into three groups depending on the age variable. The three age ranges are established as follows: Group 1 (16-27); Group 2 (28-39) and Group 3 (40 and above).

**Experience.** This is understood as the time from them becoming part of the Technical Committee of Referees to the present \((M_{\text{experience}} = 5.98; SD = 4.16)\). To organise the sample based on experience, the following four groups have been made: Group A (0-5) Group B (6-11) Group C (12-17) and Group D (18-23).

**Refereeing level.** It was decided to bring together all the referees assessed in two large groups depending on the level at which they refereed at. This separation was structured in the following way:
o Group I: Amateurs: Grass roots football and regional divisions.
A total of 97 referees belong to this group with average age ($M_{\text{age}} = 26.66; SD = 7.75$). Regarding experience ($M_{\text{experience}} = 5.48; SD = 4.14$).

o Group II: Semi-professionals: Preferente Division, Third Division and 2B Division.
The two groups belonging to the semi-professional category were made up of 24 referees ($M_{\text{age}} = 27.88, SD = 5.53; M_{\text{experience}} = 8, SD = 3.67$). Of the results related to age and experience of both groups, it can be observed that the average age and average experience is greater in the case of group 2. This can be seen as logical as it involves referees who perform their work at a higher level.

**Depending variables**

i. Response Time (RT); ii. Movement Time (MT); and iii. Average score obtained by each referee for the performance during the normal season.

**Results**

Regarding the results, the most relevant descriptive values refer to the average and standard deviation of the RT ($M = 428.13$ ms, $SD = 6.42$ ms) and of the MT ($M = 154.57$ ms, $SD = 4.21$ ms). The normality of the distribution and the majority of the variances of these two dependent variables were analysed using the Kolmogorov-Smirnov test. This test gave significant results in both variables ($p_{\text{RT}} = .003$ and $p_{\text{MT}} = .031$), so it was decided to use nonparametric statistics.

The nonparametric comparison analysis was performed for independent samples of the Mann-Whitney U Test, not finding significant differences between the groups with regard to the amateur and semi-professional levels, neither in the RT ($p = .935$) nor in the MT ($p = .132$). Subsequently, the correlation analysis was performed between the RT and MT using the Spearman correlation coefficient, finding statistically significant positive relationships ($r = .259, p < .01$).

Analysing the RT and MT variables depending on the referees’ experience in years, the results obtained through the comparison of the 4 groups established via the Kruskall-Wallis statistic for independent samples, were not significant ($p_{\text{RT}} = .633; p_{\text{MT}} = .537$). On the contrary, performing the analysis depending on the age with the same test being performed, there are significant differences between groups 2 (28-39) and 3 (40 and above), results with a different level of significance which is detailed below in the comparison of pairs using the Mann-Whitney U Test:

- Group 1 (16-27) and Group 2 (28-39): There are no differences ($p = .380$).
- Group 2 (28-39) and Group 3 (40 and above): There are significant differences ($p = .012$). In Group 3, the group with the oldest referees, they have a greater reaction time when faced with the appearance of the stimuli.
- c) Group 1 (16-27) and Group 3 (40 and above): As with the case above, there are significant differences ($p = .039$), with a greater reaction time in Group 3 than in Group 1.

The results obtained lead us to the conclusion that group 3 has a greater reaction time with regard to the appearance of stimuli, or to put it another way, the lower age groups (Groups 1 and 2) show lower reaction times to the appearance of stimuli.

With regard to the analysis on the influence of scores obtained in RT and MT and the performance score of the referee in the evaluation by the CTA, the results, calculated using the Spearman Correlation Coefficient, taking the Box Cox optimal transformation parameters, show significant correlations with RAR ($r = -.271, p = .020$); as well as the distribution of the MT with regard to the RAR ($r = -.241, p = .038$).
Discussion

The reaction time and movement time are not variables which distinguish the level referees are working at, as no statistically significant differences were found among them in either of the two variables. This could be due to there being many factors which influence reaction time, both factors which depend on the subject as independent ones (those related to the stimulus). In addition, there are those which correspond to different system of perception and development, and the response system (Pérez et al., 2011). So, it is possible that the physical condition of the subject, the starting tiredness and/or rest, motivation, as well as the body part which performs the response –dominating hand– (Henry & Rogers, 1960; Sage, 1977) can have a greater influence than the variable of the level of the referee. Nor are there significant differences with regard to the experience of the referees. However, where there have been significant differences regarding reaction time and movement time is in the age variable. Here, the results of the sample show that the older referees are, the greater their reaction time. This could have a negative influence on the speed of the referee in question when it comes to assessing a certain action during the course of the game and the time for making the correct corresponding decision.

Nonetheless, the results of this study also have a series of limitations which should be mentioned. On the pitch, the position on the field plays an important role in decision-making which is a variable which we cannot take into account in this research. The closer the referee is to the play, the lower the percentage of error. In action which takes place further away, such as offside, the referee’s assistant also has an important role in the decision-making. Referees usually offset decisions made when they have made mistakes in their perception of the game. This has a specific cost, such as letting a foul go, letting play go on, etc.

In the referee selection process, the results of the physical tests were considered, where differences in reaction time have been seen. In the future, more details will be published on this topic. These differences show us that the selection should be made more from a psychological point of view than a physical one, given that the referee who commits least errors should be on the pitch, not the referee in best physical condition.

The effect of the level of refereeing was not as important as expected. We cannot provide data which clarifies this result, nor a sufficiently convincing explanation regarding this matter. A speculative reason to justify this may be that the referees who are working at a higher level generally are of a higher age than those in the lower leagues. Considering that the reaction time is a variable which assumes worse results with the age of the subject (as has been evidenced in this study), the reaction time does not allow us to differentiate the level of the referee.

On the other hand, the results discovered through the correlational analysis between the points in the RT test and the performance score given by CTA observers are of great interest. Although the complexity and limitations the system may have been considered, given the subjectivity of the observer and the lack of precision in the categories being observed, as it is a general evaluation of performance, it does seem that referees with better scores are those which obtain better RT scores. This emphasises that the speed in which the referee indicates the foul on the pitch is being valued in a positive way. This, in turn, shows the effectiveness of the assessment process carried out using the computerised VTS system for precisely establishing this parameter.

The depth of the selection processes and promotion of football referees should be reviewed, analysing the psycho-technical tests which are performed, given that it is proposed that reaction time is a critical value in the decision-making process as found in scientific literature detailed in a systematic way (Araujo & Davids, 2015; Robles, 2014).

The psycho-technical assessment is considered essential for the objective value it gives in the intellectual behaviour of the subjects involved, who in this case are referees. This type of test is free of the usual self-report bias, meaning it is recommendable to continue using it in the selection and development of refereeing talent. From the Technical Committee of Referees of the Football Federation of Madrid, and from its Modernisation Centre,
more than seven research projects are being undertaken, linked to the Autonomous University of Madrid (UAM in Spanish), the European University of Madrid (UEM in Spanish) and the Carlos III University of Madrid. This makes it an international centre of reference for refereeing talent on a psychological and decision-making level, with studies being coordinated with FIFA and UEFA.

References


