

Development and validation of a notational system to study the offensive process in football

Hugo Sarmento^{1,2}, Teresa Anguera^{2,3}, Jorge Campaniço^{1,2}, José Leitão^{1,2}

¹Department of Sport Sciences, Exercise & Health, University of Trás-os-Montes e Alto Douro, Vila Real, Portugal,

²Centre for Research in Sport, Health and Human Movement, Vila Real, Portugal,

³Department of Methodology of the Behavioural Sciences, Universidad de Barcelona, Barcelona, Spain

Key words: match analysis; football; T-patterns; observational methodology.

Summary. The most striking change within football development is the application of science to its problems and in particular the use of increasingly sophisticated technology that, supported by scientific data, allows us to establish a “code of reading” the reality of the game.

Therefore, this study describes the process of the development and validation of an ad hoc system of categorization, which allows the different methods of offensive game in football and the interaction to be analyzed.

Therefore, through an exploratory phase of the study, we identified 10 vertebrate criteria and the respective behaviors observed for each of these criteria. We heard a panel of five experts with the purpose of a content validation. The resulting instrument is characterized by a combination of field formats and systems of categories.

The reliability of the instrument was calculated by the intraobserver agreement, and values above 0.95 for all criteria were achieved.

Two FC Barcelona games were coded and analyzed, which allowed the detection of various T-patterns.

The results show that the instrument serves the purpose for which it was developed and can provide important information for the understanding of game interaction in football.

Introduction

The study of the game by observing the behavior of the teams and players is not a recent phenomenon, characterized by a process that has evolved over time (1–3) and that has accompanied the enormous growth of sports performance for the past 50 years (4).

Traditionally, methods of analysis have used the frequency of occurrence of events (e.g., number of passes made in a certain area of the field or how many times a team committed an error) as an indicator of performance. This analysis based on the analysis of the frequency of certain performance parameters provided and continues to provide important information for coaches and athletes, enabling advances in training processes (5).

However, the game of football is characterized by great complexity of constraints that make it difficult to objectify its observation and analysis. Therefore, in order to achieve a greater similarity to the objectification of the actions observed in this type of games, underpinned by the “bond” or invariances on the one hand, and variations or by chaos¹ and

disorder (i.e., the randomness and variability that characterize the game), on the other hand, the strict delimitation of models to guide the action of the observer gains fundamental importance. These models cannot be confined solely to the mere registration of the occurrence of events.

Although the tactical constraint plays an important role in the performance of football teams, the history of the scientific analysis of this game has been driven by biomechanical and physiological approaches, paying little attention to the tactical behavior of the players and team organization. For coaches and researchers, tactical analysis can be helpful, since they offer the opportunity to identify match regularities and random features of game events (6).

The scientific analysis of sports performance aims to understand the behavior in the game in or-

¹Chaos Theory attempts to explain the phenomena of characteristics of order and stability on the one hand and accompanied by disorder and irregularity on the other. It refers to the origin and nature of matching patterns of uniformity and variety in the behavior of systems (6).

der to be able to improve their quality (7). It seems clear that the behavior will only become a subject of scientific research if its registration and quantification are carried out in an objective and rigorous way (8). It is important to say that in many situations, the only way to study the behavior without excessively distorting it is watching it as it occurs spontaneously.

However, much of the research conducted in the area of sports performance tends to observe the action and to describe the behavior of players who are in possession of the ball, excluded from this analysis the behavior of other players, more specifically the context in which the action occurs (7).

Moreover, the methods used to collect data in some studies are also questionable, placing the problem of its reliability. It is here that we stand before one of the key issues of observation, the compatibility “technical observation” versus “observational methodology” (9).

In this field of research (match analysis), there is a growing need for rigorous analysis of data, particularly in terms of identifying the sources of variability resulting from methodological issues, for example, those caused by varying methods of data collection used in studies, the lack of standardized approaches, and especially, the considerable differences in the approaches used to classify motion (10).

The structure of the football game is extremely complex, both in terms of formality and functionality. Underlying an apparently simple game is a high motor, strategic and social complexity, which should be taken into account in its study and analysis (11).

According to McGarry (7), the challenge is to obtain more appropriate scientific description of game behavior that can be quantified objectively, with more richness of details than that obtained hitherto, the aim being a complete system description that approximates the vast amount of contextual information that is readily available from visual inspection.

Therefore, in order to achieve deeper insight into the football team tactical game, it is necessary to record the substantial tactical actions in a chronological, sequential order, so the stream of tactical behavior can be recognized (12).

Given the perspective that the observational methodology gives us, all team sports, and football in particular, allow a broad range of possibilities for study. From the analysis of success factors of the relationship among game players, and a study of the team’s social behavior until the analysis of the sequences of the driving behaviors of certain players (13).

Anguera (14) considers, in this context, that the observation in sports requires a clear definition of its scope of activity, particularly in two areas: content

and process or methodology. It is here that the Observational Methodology exposes all its potential, since the multivariate characteristics of behaviors in a competitive situation, the interaction that occurs among them and the difficulty in controlling the contextual variables, never identical in different situations, recommend the use of this methodology. From the above-mentioned point of view, this methodology opened the doors to the description and analysis of social-driven dynamics from the logic of scientific research (15).

If performance analysis is to continue to advance understanding of sports performance, then it must continue to explore better methods of collecting and analyzing data. In this paper, we introduce and explain a new data analysis method that has the potential to make a significant contribution to analyses of sports performance (16).

The collection and data analysis will enable the analysis of several tactical-technical indicators and their influence on outcome. Among these techniques we highlight: i) analysis of T-patterns, which allows the detection of hidden patterns of behavior, ii) the sequential analysis, which allows detection of behavior patterns, as well as the demand for significant association relationship between behaviors recorded during these sequences, iii) analysis of polar coordinates, which enables a vector representation of the complex network of interrelationships established among the various groups making up the observational method to be used.

Among the above-mentioned techniques, we highlight the T-pattern detection, which allows the detection of the temporal and sequential structure of a data set. The method has been developed, outside the sports, on the assumption that complex streams of human behavior have a temporal/sequential structure than cannot be fully detected through unaided observation or with the help of standard statistical and behavior analysis methods. Given that observational records of human behavior, including sport performance analysis, have both a temporal and sequential structure, an analysis tool that can describe this structure will enhance understanding of the behaviors being studied (17, 18).

Given that the observational methodology is one of the options of the scientific study of human behavior (9), the aim of this study was to describe the stages of construction and validation of an observation system in football, to reliably register behaviors and interactions of the players of high-performance teams.

Methods

This study describes the development of a specific notational system to study the offensive process (OP) in football, specifically to study counter-

Table 1. Vertebrates criteria of the instrument

	Definition
Characterization of the game	Spanish, English and Italian Leagues 1°/2° part Home/away Momentary result
Offensive game methods	Counter attack Fast attack Positional attack
Start of the offensive process	Recovering possession (RP) of the ball by interception RP of the ball by disarm RP of the ball by goalkeeper action RP of the ball by interruption due to the rules of the game RP of the ball by opponents goal
Development of the offensive process	Development (D) by short/medium pass D by long pass D by conduction of the ball D by reception/control D by dribble (1×1) D by duel D by goalkeeper action of the team that is in the offensive process D by shoot D by cross D by unsuccessful opponent intervention D by opponent goalkeeper action
End of the offensive process	Shoot with score goal Inside shot Defended shot by goalkeeper Shot out Shot against opponent Free kick Corner kick Penalty Pass inside the penalty area Recovery of the ball by the opponent Pass to the outside of the field Violation of the rules of the game
Way and direction of the pass	Pass to the front Pass to the back Pass to the side Diagonal pass to the front Diagonal pass to the back
Height of the pass	Shallow pass Medium high pass High Pass
Rhythm of the game	Fast rhythm of the game Slow rhythm of the game
Spatial characterization	Zone 1 to zone 12
Game Center	Relative numeric inferiority Absolute numeric inferiority Equality numeric under pressure Equality numeric unpressured Relative numeric superiority Absolute numeric superiority

attack, fast attack, and positional attack. This system follows on from other studies carried out in this area of research (19, 20) and results from an adaptation to the specific objective of this study.

Procedure. In order to prepare the observational instrument, in addition to literature review, a panel of experts consisting of graduate students, researchers, and experienced coaches were heard to find out what performance indicators should be included in the study.

Based on previous procedures and an exploratory phase of the study, in which we observed video images of several offensive sequences in 9 games of the Spanish, Italian and English Leagues, we defined the vertebrate criteria of the instrument and proceeded to draw up a list of behaviors/situations observed for each of the criteria (Table 1).

Then, through an exploratory phase, an open list of observed behaviors/situations, corresponding to each of the criteria was produced.

Therefore, this construction converges in a system where all behaviors were likely to register and that each one only corresponds to a single category. A list called the catalog was then drafted, for each of the criteria. In Table 2, it is possible to check all behaviors assigned to the criterion of "way and direction of the pass."

The list of criteria and behaviors produced was then presented to a panel of 5 experts for the purpose of content validation. These experts met both the following conditions: i) coach a football team for

Table 2. Conceptual definition and categories for the criteria "Way and direction of the pass"

Way and direction of the pass	
Represents the direction of the ball during its movement between two elements of the same team with the aim of continuing the game to allow the continued and successful possession of the ball. We propose 5 possible categories to guide the action techniques – Way/Direction.	
Categories	Definition
Pass to the front	Whenever the player in possession performs a pass towards the opponent goal.
Pass to the back	Whenever the player in possession performs a pass towards the defended goal.
Pass to the side	Whenever The player in possession performs a lateral pass to the axis of attack.
Diagonal pass forward	Whenever the player in possession performs a pass diagonally to the axis of attack towards the opposing goal (the ball moves to the corridor and sector).
Diagonal pass back-wards	Whenever the player in possession performs a pass diagonally to the axis of attack towards the goal defended (the ball moves to the corridor and sector).

over five years; ii) have conducted research in this area of study. This phase resulted in minor changes resulting from appropriate suggestions.

After this stage, we prepared a macro in Excel, which allows us to collect the data more quickly, efficiently, and with less likelihood of mistakes. The page has been programmed specifically for data to be collected in accordance with the requirements for the software THEME.

Fig. 1 represents the initial window of this macro. At this stage, criteria for the following must be reported: i) the context of the game (league, part (first vs. second), place (home vs. away), momentary result, numeric relationship); ii) for the offensive game method (fast attack, positional attack, counter-attack); iii) for the start of the offensive process (behavior for the ball recovery, context of interaction, and field area); iv) and for the exact time of the start of the offensive sequence. The configuration of the page only allows one to skip to the next window, when all fields have been completed. The next window allows the coding of all procedures related to the development of the offensive process, and so on, until the last window, which allows the coding of the end of the offensive process.

Fig. 2 shows an example of data derived from the coding sequence of a counter-attack. This sequence has developed into a stage of the game where the teams were tie at the marker (G0). It refers to a counter-attack, which occurred in a situation where the teams were playing with the same number of elements (IN, CA). The start of the offensive process took place through the recovery of possession of the ball by interception, in zone 9 of the field, in a context of interaction that is characterized by unpressured equality (Ipi, Z9, SPinp). Then, the development of the offensive process was achieved by a short pass, in zone 8 of the field, shallow pass to the front, allowing development through a high rhythm of the game, in a context of relative numeric superiority (Dpc, Z8, PFr, Pr, Rjr, SPsr). The offensive process had a few more sequences of development until it finished (end of the offensive process) through a player who shoots against the opponent, in zone 3 of the field, in a context of the interaction characterized by relative numeric inferiority (Fca, Z3, Pir).

Intrarater reliability. In studies of this scope, the stability of observations has a significant importance, i.e. the reliability of data collection. In this study, we chose to analyze the data quality through the intra-observer agreement, which was verified through the kappa reliability index.

Two FC Barcelona games, in the 2009/2010 sporting season, were randomly selected to be encoded twice. These observations were made with an interval of at least two weeks to minimize the difference between ratings.

Fig. 1. The first window of the macro in Excel

	A	B	C
1	Dataname	Frames	Events
2	Llg1pC	0	:
3	Llg1pC	35350	G0
4	Llg1pC	35351	IN,CA
5	Llg1pC	35352	Ipi,Z9,SPinp
6	Llg1pC	35375	Dpc,Z8,PFr,Pr,Rjr,SPsr
7	Llg1pC	35400	Dcd,Z9,Pr,Rjr,SPsr
8	Llg1pC	35425	Dpc,Z4,PPI,Pr,Rjr,Pir
9	Llg1pC	35450	Drc,Z4,Rjr,Pir
10	Llg1pC	35475	Dcz,Z1,Rjr,Pir
11	Llg1pC	35525	Fca,Z3,Pir
12	Llg1pC	35526	Y

Fig. 2. Codification of an offensive sequence

Software. For the analysis of data quality, we used the software GSEQ-SDIS (21) and its function, calculate kappa. For the detection of temporal patterns, we used the software THEME.

Results

The observation of these two games allowed the coding of 42 offensive sequences in each observation (Table 3).

Table 3. Matches recorded

Place	Result	Teams	Data rows in record file
Home	Won	Mallorca	310
Home	Won	Real Madrid	353

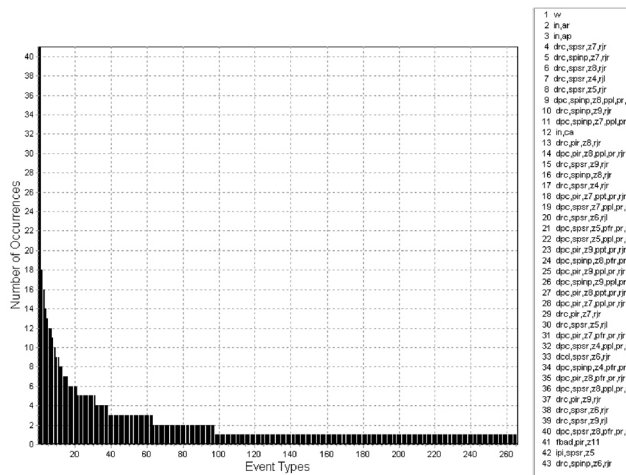


Fig. 3. Event frequency chart

The software also enables the analysis of the frequency of coded events (Fig. 3). Through this analysis, we concluded that the offensive game method more often used by the FC Barcelona in these two games was fast attack, which was followed by positional attack. Counter-attack was the least used method of attacking.

Intrarater reliability. The analysis of the quality of data, resulting in the following indices of agreement of Cohen's kappa: start of the offensive process (0.98); development of the offensive process (0.97); and end of the offensive process (0.93). By examining the values obtained for each criterion, we concluded that the results showed a high stability be-

tween observations, because all had a kappa value more than 0.95.

T-patterns. The analysis of the transcribed records revealed a large number of complex structures that are repeated over time in these two observed games. For pattern detection, the following criteria were used: the minimum number of events was set at 3 and the level of significance was set at 0.05.

As seen in previous studies in this area of research (5, 10), the data show that football has a large number of temporal patterns. As in the study carried out by Jonsson et al. (2006), the number, frequency, and complexity of the patterns found show that the behavior in sports presents a more synchronized structure than that which the "human eye can detect."

The pattern in Fig. 4 occurred twice during the first half of the match Barcelona vs. Real Madrid.

In this case, the two teams to play the game were in a position of numerical equality (11×11), and the momentary result of the game was also equal.

The beginning of the offensive process is done by recovery possession of the ball by an interception, in zone 5 of the field, in the context of relative numeric superiority (3×2). The development of the offensive process occurs by performing a shallow pass to the front, in zone 5 of the field, which resulted in a high rhythm of the game. This sequence finished without effectiveness, by the recovery of possession of the ball by an opponent in zone 11 of the field.

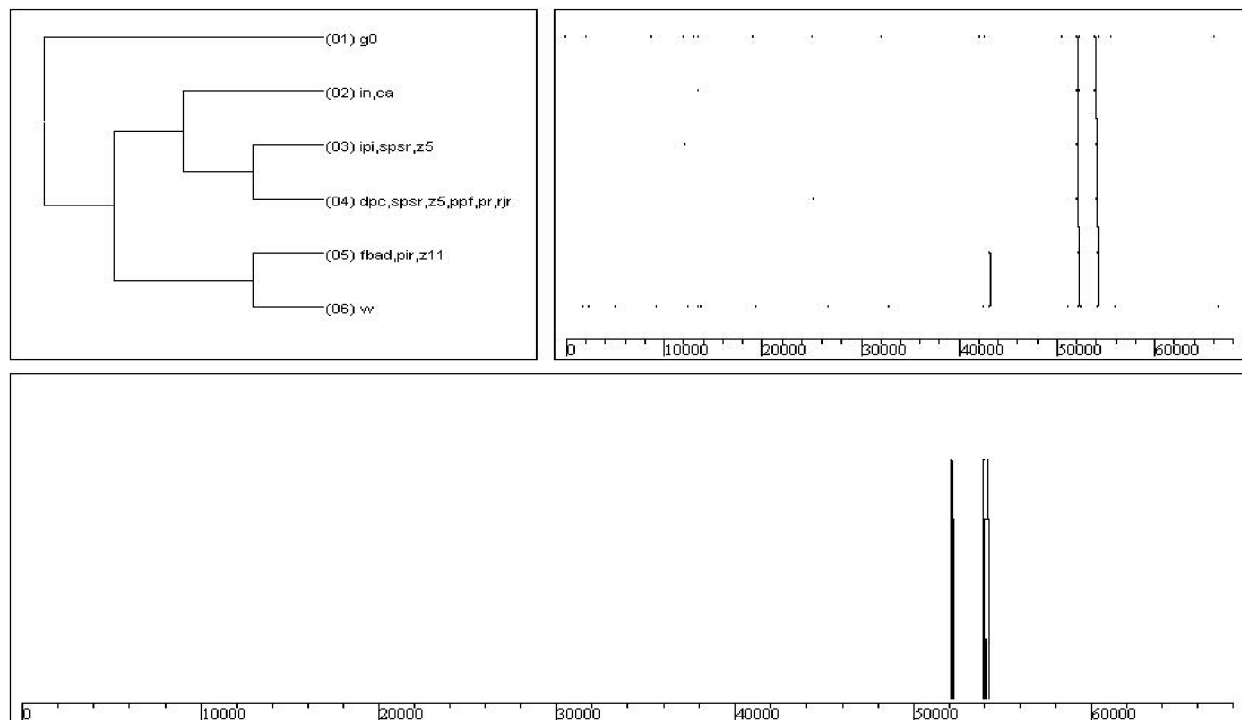


Fig. 4. An example of a detected T-pattern

Discussion

The main subject of tactical analysis should not be the players' actions, taken disjointedly, but the sequences of game play resulting from the actions that occur during the different phases of the match (6).

Therefore, the construction of ad hoc instruments to enable the collection of data through the use of systematic observation can significantly increase the capacity for understanding and analyzing the context in which the behaviors occur in football players.

In football, the high number of players, the complexity of tactical behaviors, and the speed at which the actions are generated inhibit the observational record of driving behaviors of interaction.

However, the observational instrument presented in this study and the results achieved in relation to data quality have been completely satisfactory.

Unlike the studies that focus exclusively on the frequency analysis, this instrument allows us to analyze the offensive process in football from a qualitative perspective. The results show that many temporal patterns exist in football. The number, frequency, and complexity of the detected patterns indicate that sports behavior is more synchronized than the human eye can detect.

Conclusions

The study of the offensive game methods from this perspective allows the identification of hidden patterns of behavior (T-patterns) that are not identifiable through a simple observation of the "naked eye."

The preliminary results show the potential of the proposed study design, allowing us to go beyond traditional analysis of frequency of sports behaviors, which may help all those involved in the training process to understand the sporting performance more deeply.

The results suggest that collecting a larger volume of data allowed us to detect more temporal patterns of behavior. The analysis of these patterns, complemented with information derived from basic statistics and with the understanding that the coaches have in relation to this structured behavior, may contribute significantly to optimize sports performance.

Acknowledgments

We gratefully acknowledge the support of the Portuguese Foundation for Science and Technology (SFRH/BD/45736/2008).

Nacionalinės sistemos futbolo atakos procesams tirti sukūrimas ir pritaikymas

Hugo Sarmento^{1,2}, Teresa Anguera^{2,3}, Jorge Campaniço^{1,2}, José Leitão^{1,2}

¹Trás-os-Montes e Alto Douro universiteto Sporto mokslų, fizinio lavinimo ir sveikatos katedra, Vila Real, Portugalija,

²Sporto, sveikatos ir žmogaus judėjimo tyrimų centras, Vila Real, Portugalija,

³Barselonos universiteto Elgesio mokslų metodologijos katedra, Barselona, Ispanija

Raktažodžiai: varžybų analizė, futbolas, T vaizdinys, stebėjimo metodologija.

Santrauka. Reikšmingiausias pokytis futbolo sporto raidoje yra mokslo pritaikymas jo problemoms spręsti, ypač besiplečiantis patikimų technologijų panaudojimas, kurios pagrįstos moksliniais duomenimis, leidžiantis sudaryti „nuskaitymo kodus“ atkuriant žaidimo realijas.

Pateiktas tyrimas aprašo sukūrimo ir pritaikymo procesus kuriant kategorijų sistemą, pagal kurią galima nagrinėti skirtingas puolimo metodikas žaidžiant futbolą.

Metodas. Tiriamosios fazės metu nustatėme 10 esminių kriterijų ir jiems priskyrėme stebimą atitinkamą situaciją, elgseną. Sudarėme penkių ekspertų grupę, siekdami patvirtinti jų turinį. Galinis instrumentas nusakomas deriniu – lauko forma ir kategorijų sistema.

Rezultatai. Instrumento analizės patikimumas apskaičiuotas pagal stebėtojų sprendimų suderinamumą, patikimumas daugiau nei 0,95 buvo pasiektas visiems kriterijams.

Aprašyti ir išanalizuoti du Barselonos klubų žaidimai. Remiantis šia analize, galima nustatyti įvairius T vaizdinius.

Išvada. Rezultatai parodė, jog instrumentas efektyvus, siekiant užsibrėžto tikslo.

References

1. Hughes M, editor. Notational analysis of soccer. 5th International Scientific Conference on Kinesiology; 2008 10-14 September; Zagreb: Faculty of Kinesiology, University of Zagreb, Croatia.
2. Hughes M. Performance analysis – a 2004 perspective. Int J Perfor Anal Sport 2004;4:103-9.
3. Garganta J. A análise da performance nos jogos desportivos – Revisão acerca da análise do jogo. Rev Portug Ciências Desporto 2001;1(1):57-64.
4. Kunh W, editor. Changes in professional soccer: a qualita-

- tive and quantitative study Science and football V; Lisbon: Routledge; 2005.
5. Borrie A, Jonsson GK, Magnusson MS. Temporal pattern analysis and its applicability in sport: an explanation and exemplar data. *J Sports Scie* 2002;20(10):845-52.
 6. Garganta J. Trends of tactical performance analysis in team sports: bridging the gap between research, training and competition. *Revista Portuguesa de Ciências do Desporto*. 2009;9(1):81-9.
 7. McGarry T. Applied and theoretical perspectives of performance analysis in sport: scientific issues and challenges. *Int J Perfor Anal Sport* 2009;9:128-40.
 8. Leitão J, Campaniço J. Research methods support in observation sports laboratory. *Motricid Sport Health Hum Develop* 2009;5(3):27-33.
 9. Anguera T, Villaseñor Á, Lopes LL, Mendo AH. La metodología observacional en el deporte: conceptos básicos. *Lecturas: Educación Física y Deportes* [serial on the Internet]. 2000;(24). Available from: URL: <http://www.efdeportes.com/efd24b/obs.htm>
 10. Jonsson GK, Anguera MT, Blanco-Villasenor A, Losada JL, Hernandez-Mendo A, Arda T, et al. Hidden patterns of play interaction in soccer using SOF-CODER. *Behav Res Methods* 2006;38(3):372-81.
 11. Anguera T. Metodología básica de observación en fútbol. In: Paidotribo, editor. *Metodología de la enseñanza del fútbol*. Barcelona; 2003. p. 303-21.
 12. Pfeiffer M, Perl J. Analysis of tactical structures in team handball by means of artificial neural networks. *Int J Comp Sci Sport* 2006;5(1):4-14.
 13. Anguera T. Hacia la búsqueda de estructuras regulares en la observación del fútbol: detección de patrones temporales. *Cult Ciencia Deporte* 2004;1(1):15-20.
 14. Anguera T. Methodological observation in sport: Current situation and challenges for the next future. *Motricid Sport Health Hum Develop* 2009;5(3):15-25.
 15. Mendo H, Anguera T, Santos R. Desarrollo y optimización de una herramienta observacional en el tenis de individuales. *Psicothema* 2005;17(1):123-7.
 16. Gudberg J, Anguera T, Sánchez-Algarra P, Oliveira C, Campaniço J, Castañer M, et al. Application of T-pattern detection and analysis in sports research. *Op Sport Sci J* 2010;3:95-104.
 17. Magnusson M. Hidden real-time patterns in intra and inter-individual behaviour: description and detection. *Eur J Psychol Assess* 1996;12(112-23).
 18. Magnusson M. Discovering hidden time patterns in behaviour: T-patterns and their detection. *Behav Res Methods Instrum Comput* 2000;32:93-110.
 19. Barreira D. Transição defesa ataque em futebol. *Análise sequencial de padrões de jogo relativos ao campeonato português 2004/2005*. Porto: Universidade do Porto; 2006.
 20. Lopes J. *Análise diacrónica heterocontingente dos métodos de jogo ofensivo no futebol*. Porto: Universidade do Porto; 2007.
 21. Bakeman R, Quera V, editors. *Análisis de la interacción. Análisis secuencial con SDIS y GSEQ*. Madrid: Ra-Ma; 1996.

Received 17 May 2010, accepted 7 June 2010

Straipsnis gautas 2010 05 17, priimtas 2010 06 07