

## Application of logistic regression models in observational methodology: game formats in grassroots football in initiation into football

Daniel Lapresa<sup>\*1</sup>, Javier Arana<sup>2</sup>, M. Teresa Anguera<sup>3</sup>, J. Ignacio Pérez-Castellanos<sup>1</sup> y Mario Amatria<sup>1</sup>

<sup>1</sup> Universidad de La Rioja (Spain).

<sup>2</sup> Universidad Internacional de La Rioja (Spain).

<sup>3</sup> Universidad de Barcelona (Spain).

**Título:** Aplicación de modelos de regresión logística en metodología observacional: modalidades de competición en la iniciación al fútbol.

**Resumen:** El presente trabajo constituye un ejemplo de las posibilidades de utilización de modelos de regresión logística -múltiple y simple- en el ámbito de la metodología observacional y más concretamente en su aplicación a la actividad física y el deporte. La ejemplificación se concreta en un estudio que pretende discriminar qué modalidad de juego (fútbol 3 o fútbol 5) constituye una propuesta más adaptada a las posibilidades del niño de 6 a 8 años. La satisfacción de ambos objetivos -metodológico y disciplinar- se satisface mediante un modelo de regresión logística múltiple, que va a analizar en qué modalidad el niño obtiene un desempeño espacial de mayor calidad; y tres modelos de regresión logística simple, que van a aportar información relevante acerca de en qué modalidad de fútbol el niño obtiene un desempeño técnico-táctico más eficaz y adaptado.

**Palabras clave:** metodología observacional; regresión logística; fútbol; iniciación.

**Abstract:** This study shows how simple and multiple logistic regression can be used in observational methodology and more specifically, in the fields of physical activity and sport. We demonstrate this in a study designed to determine whether three-a-side futsal or five-a-side futsal is more suited to the needs and potential of children aged 6-to-8 years. We constructed a multiple logistic regression model to analyze use of space (depth of play) and three simple logistic regression models to determine which game format is more likely to potentiate effective technical and tactical performance.

**Key words:** observational methodology; logistic regression; football; grassroots.

### Introduction

Regression is a statistical technique used to analyze functional relationships between variables (Silva & Barroso, 2004). Regression models are used to evaluate the association between a given variable, called the criterion variable (or dependent in experimental studies), and one or more predictor variables (or independent). They can be used for estimation or prediction purposes. In the first case, they examine the extent to which the criterion variable is explained by the predictor variable(s) and in the second case, they are used to predict a criterion variable based on values obtained for the predictor variable(s) (Hellewicz, 2007; Herrera & Gómez, 2008). A range of regression models exist that vary in numerous aspects, including number of predictor variables (simple vs multiple regression) and the nature of the variables (continuous or dichotomous).

The logistic regression model (Ato & López, 1996) uses a dichotomous criterion variable and one or more qualitative, ordinal, or quantitative predictor variables. Logistic regression analysis is mainly used in health sciences (Bagley, White & Golomb, 2001; de Irala, 1999; Steyerberg, Borsboom, Van Houwelingen, Eijkemans & Habbema, 2004; Steyerberg, Eijkemans, Van Houwelingen, Lee & Habbema, 2000), including studies of physical activity and sport (Ábalo, 2012; Araguay, Mena & Vavarello, 2004; Fernández, de la Cruz Márquez, Cueto, de la Cruz Campos & Salazar, 2008), and in behavioral sciences (Broc & Gil, 2008; Rando, Blanca

& Frutos, 2000; Rodrigo, Molina, García & Pérez, 2012). In recent years, several studies employing observational methodology have successfully used logistic regression analysis applied to sport (Arana, 2011; Arana, Lapresa, Anguera & Garzón, 2013; Ardá, Maneiro, Rial, Losada & Casal, 2014; Casal, 2010).

Logistic regression is one of the analytical techniques proposed by Anguera, Blanco-Villaseñor, Hernández-Mendo and Losada (2011) for studies employing an observational design. The present study had two aims: to show how logistic regression analysis could be applied to an observational study in the field of physical activity and sport and to determine how futsal competition models could be optimized for children aged 6-8 years old. The design fits into new research designs that are emerging in the field of sports psychology (Ortín-Montero, Moya-Faz & Garcés de los Fayos, 2013). As highlighted by Romero & Vegas (2003), five-a-side futsal (F-5) is the smallest small-sized game format for children in this sport, even though it has been shown that children aged 6-8 have difficulties in demonstrating competence in this format (Lapresa, Arana & Carazo, 2005; Torres & Rivera, 1994; Wein, 1995). F-5 has essentially the same format (same pitch and same ball) as adult futsal (Lapresa, Álvarez, Arana, Garzón & Caballero, 2013). In this study, we compared F-5 with an alternative format (3-sided futsal, or F-3) to see if the latter might be more suited to the needs and potential of children aged 6-8 years old. We specifically compared the use of space (depth of play) and ball handling skills (effectiveness and suitability for age) during offensive play in F-3 and F-5.

**\* Dirección para correspondencia [Correspondence address]:**

Daniel Lapresa Ajamil. Universidad de la Rioja. Departamento de Ciencias de la Educación. Edificio Vives, C/ Luis de Ulloa s/n. 26004. Logroño, La Rioja (España). E-mail: [daniel.lapresa@unirioja.es](mailto:daniel.lapresa@unirioja.es)

## Method

We employed observational methodology (Anguera, 1979) with the following design, based on the work of Anguera, Blanco-Villaseñor and Losada (2001) and Anguera et al. (2011): point (i.e. the aim was not to compare teams or follow them over time); within-session (continuous observation of behavior recorded during a single session (Anguera & Hernández-Mendo, 2013); nomothetic (separate teams not acting as a unit); and multidimensional (multiple criteria in the observation instrument). Observation was non-participatory and governed by scientific criteria, and the level of perceptibility was complete.

### Participants

To analyze the performance of 6-to-8-year-olds in F-3 and F-5, we used a convenience sample (Anguera, Arnau, Ato, Martínez-Arias, Pascual & Vallejo, 1995) selected from the top three teams in the F-5 league organized by the Football Federation in the Spanish autonomous community of La Rioja. To collect the data for the study, we organized a three-way tournament for each of the game formats. To ensure between-session consistency, we used the same pitch, the same game duration, the same referee, the same ball (no. 5), and the same division of the pitch. F-3 was played according to the rules established by the Royal Spanish Football Association in Lapresa, Arana, Garzón, Egüén and Amatria (2008).

### Instruments

#### a) Observation Instrument

We used the observation instrument designed by Lapresa, Arana, Garzón, Egüén and Amatria (2010) to study technical actions executed by young football players. The observation structure was the same as that used in the different versions of the Football Observation System (Anguera et al. 2003).

The instrument comprised a combination of field formats and category systems (Anguera, Magnusson & Jonsson, 2007) that met the needs of our observational design. This combined approach is feasible when the field format criteria, used to build the category system, are atemporal and supported by a theoretical framework. Table 1 summarizes the criteria applied in the observation instrument.

**Table 1.** Observation Instrument.

No.	Criterion	Type of criterion	Brief description: codes
1	Possession of ball	Fixed	PO) Possession of ball by team being observed; PC) Possession of ball by opposing team; Inob) Unobservable
2	Play no.	Fixed	1, 2, 3... n
3	Move initiation zone	Variable	ZS10, ZS20, ZS30 -safety sector-; ZS40, ZS50, ZS60 -creation sector in own half-; ZS41, ZS51, ZS61 -creation sector in opposing half-; ZS70, ZS80, ZS90 -definition sector-
4	Move conclusion zone	Variable	ZE10, ZE20, ZE30 -safety sector-; ZE40, ZE50, ZE60 -creation sector in own half-; ZE41, ZE51, ZE61 -creation sector in opposing half-; ZE70, ZE80, ZE90 -definition sector-
5	Contact with ball	Variable	C1) single touch and regulatory throw-in/kick-in; C12) attempt to control the ball with 2 or more touches resulting in loss of ball; C2) control of ball -including picking up of ball by goalkeeper- followed by a shot* (*= regardless of whether the ball reaches a team member or is recovered by an opponent); C23) control of ball, followed by dribbling, and loss of ball; C24) control of ball, followed by dribbling, attempt to go around one or more opponents, and loss of ball; C3) control of ball, followed by dribbling and shot*; C4) control of ball, passing of one or more opponents, and shot*; C5) Header
6	Interruptions	Variable	FDFT) free kick for team being observed; FDSN) neutral kick; FFSB) throw-in for team being observed; FFSE) corner kick for team being observed; FFSP) goal kick for team being observed; CDFI) free kick against team being observed; CDSN) neutral kick; CFFB) throw-in against team being observed; CFFF) corner kick against team being observed; CFSP) goal kick against team being observed
7	Interceptions	Variable	P) loss of ball; R) recovery of ball; IOC) Occasional interception with continuation of play
8	Shot	Variable	TG) shot resulting in goal; TI) shot intercepted by opponent other than the goalkeeper; TM) shot between the posts not resulting in a goal; TF) shot wide of the posts; TP) shot saved or cleared by the goalkeeper

#### b) Recording instrument

The sessions were recorded using the software program ThemeCoder, with consideration of the work of Jonsson (2006) and Jonsson et al. (2006).

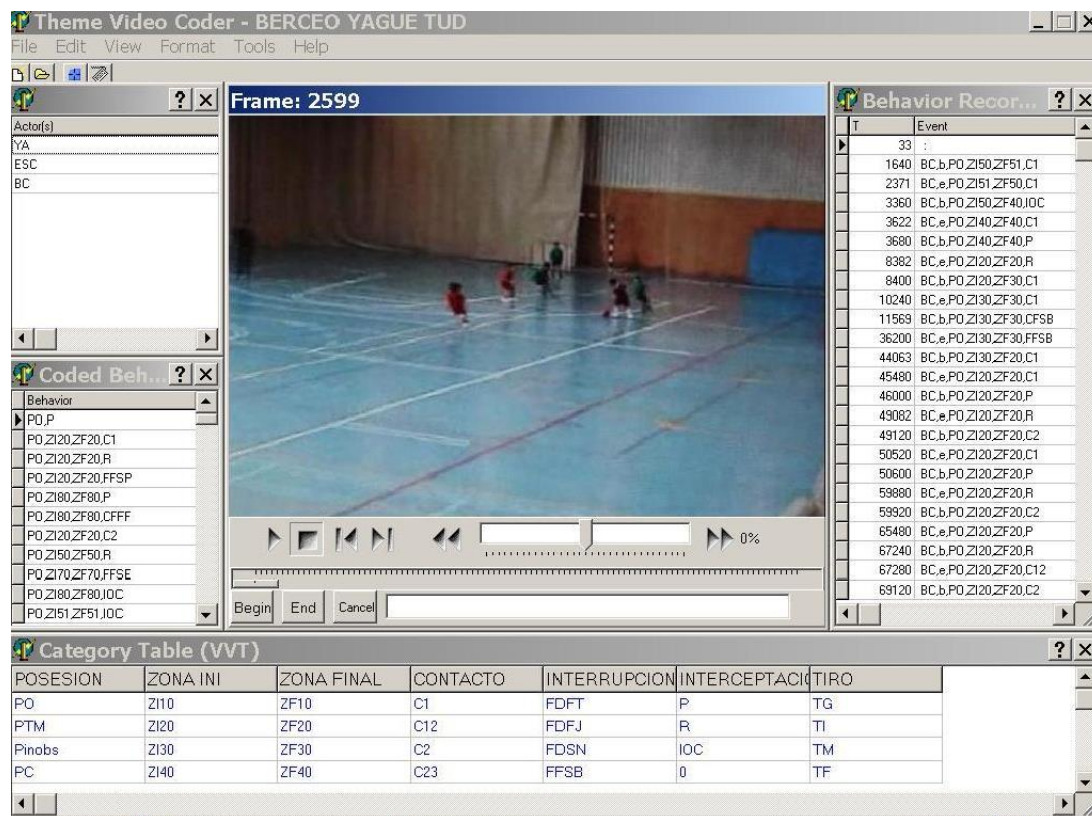


Figure 1. Screenshot showing data recording procedure in ThemeCoder.

### c) Analysis instrument

The data collected were recoded for analysis with SPSS, version 19.0.

#### Procedure

Video recordings of each of the matches were used to record the study data. Each match consisted of a given number of moves, which, in turn, consisted of a given number of actions. To study the use of space, we analyzed 582 moves for F-3 and 361 moves for F-5, and to analyze the technical performance of players, we analyzed 1361 contacts for F-3 and 881 for F-5.

We designed a multiple logistic regression model to determine the likelihood of a move being successful in F-3 and F-5 according to the use of space (depth of play). The criterion variable (Move Conclusion) was dichotomous; the two possibilities were Success (if the move concluded in Zone 80, which contained the rival team's goalposts) or Failure (if the move concluded outside this zone). Two predictor variables were considered: Move Initiation Zone (Safety Sector, Creation Sector in Own Half, and Creation Sector in Opposing Half) and Game Format (F-3 and F-5). Because the criterion variable was dichotomous and there was more than one predictor, we used a multiple logistic regression model.

Because the predictor variable Move Initiation Zone had more than two categories, we generated dichotomous dummy variables, maintaining the information provided by the original variable. Two reference categories were used: Definition Sector for Move Initiation Zone and F-5 for Game Format.

The multiple logistic regression model employed is shown below, where:  $P(Y)$  is the probability of an event occurring;  $e$ , Euler's constant  $\approx 2.718281$ ;  $X_i$  ( $i=1,2,\dots, k$ ), the predictor variables;  $\beta_0$ , the constant; and  $\beta_i$  ( $i=1,2,\dots, k$ ), the logistic regression coefficients

$$P(Y) = \frac{1}{1 + e^{-(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k)}}$$

We designed three simple logistic models to analyze the children's technical performance in F-3 and F-5. The three criterion variables were: a) Success: -contact resulting in a shot at goal or continuation of the attack- versus Failure -contact not resulting in either of a shot or continuation-; b) Adapted -contact involving dribbling- versus Not Adapted -contact did not involving dribbling-; c) Success+Adapted -contact involving dribbling resulting in a shot at goal or continuation of the attack- versus all other combinations.

As just one predictor is used in simple regression models, we used the same model as that used for the multiple analysis except that it contained just one regression coefficient:  $(\beta_0 + \beta X)$ .

## Data quality control

To assess the reliability of the data collected using the observation tool, we analyzed the level of agreement between the different data sets. Agreement was evaluated using the kappa statistic (Cohen, 1960), which is a measure of agreement for nominal classifications, with no ordering of data among the different categories. In this study, the same observer recorded, on two separate occasions, at least 10% of each of the matches that made up the study sample. Kappa was calculated using version 5.1 of the SDIS-GSEQ software program following the recommendations of Bakeman & Quera (1995, 2001, 2011). According to the criteria of Landis and Koch (1977, p. 165), the agreement between the six data sets generated for each game format (team observed/match) was "almost perfect" (F-5: .909, .906, .942, .934, .9544, .918; and F-3: .909, .906, .942, .934, .954, .918).

## Results

### Multiple logistic regression

The criterion variable Success (move ending in zone containing the rival's goal) was significantly associated with the predictors Move Initiation Zone ( $\chi^2= 46.754$ ;  $p\leq .001$ ) and Game Format ( $\chi^2= 24.687$ ;  $p\leq .001$ ) and we therefore included these predictors in the multiple logistic regression model. The association between the two predictors was not statistically significant ( $\chi^2= 4.495$ ;  $p= .213$ ), allowing us to rule out collinearity. We next analyzed the possibility of confounding and interaction. In the first case, the different Exp (B) for Move Initiation Zone exhibited a variation of less than 10% when this variable was incorporated into the Game Format variable model. It was therefore decided to maintain the three dummy categories for this predictor in the multiple logistic regression model. Interaction between the predictor variables was also ruled out, as when this interaction was introduced into the model, the Game Format variable coefficient lost its significance (Álvarez, 2007).

To choose which variables to include in the predictive regression model and to estimate the strength of the association between the criterion variable and the predictor variables, we used three SPSS procedures: Enter, Forward Selection (Wald), and Backward Elimination (Wald). The three procedures included the same variables in the logistic regression model: Safety Sector, Creation Sector in Own Half, Creation Sector in Opposing Half, and F-3.

Once the variables were selected, we evaluated the discriminatory power of the multiple logistic regression model. The model accurately predicted 70.2% of the cases (moves) analyzed. Furthermore, it yielded a sensitivity of 19.9% (ability to correctly predict a successful move) and a specificity of 91.7% (ability to correctly predict an unsuccessful move).

We then performed the estimation process based on the resulting logistic regression model. In the first case, the Exp (B) value corresponding to Game Format was 2.426 (>1 and

significant), indicating that the probability of a move ending in Zone 80 was 2.426 times higher for F-3 than for F-5 (Table 2). In the second case, the predicted probability of success for F-3 was .226 for moves that started in the Safety Sector, .344 for those that started in the Creation Sector in Own Half, and .529 for those that started in the Creation Sector in Opposing Half. The respective results for F-5 were .107, .178, and .316.

**Table 2.** Estimated results for model coefficients.

	B	SE	Wald	DF	Sig.	Exp (B)
Action Initiation Sector			50.536	3	0.000	
Safety Sector	-1.149	0.219	27.547	1	0.000	0.317
Creation Sector Own Half	-0.566	0.220	6.639	1	0.010	0.568
Creation Sector Opposing Half	0.195	0.222	0.767	1	0.381	1.215
Game Format (F-3)	0.886	0.162	29.846	1	0.000	2.426
Constant	-0.963	0.198	23.663	1	0.000	0.382
Game Format (F-5)	-0.886	0.162	29.846	1	0.000	0.412
Constant	-0.077	0.175	0.194	1	0.660	0.926

### Simple logistic regression analyses

There was a significant association between each of the criterion variables (Success, Adapted, and Success+Adapted) and Game Format ( $p\leq .05$ ). We therefore established Game Format as the criterion variable in the three simple logistic regression models, which we constructed using the Enter selection method in SPSS.

The first model had a predictive accuracy of 55.6%, a sensitivity of 66.4%, and a specificity of 44.9%. The Exp(B) value for Game Format was 1.610 (>1 and significant), indicating that the probability of a contact involving a shot or continuation of attack was 1.610 times higher for F-3 than for F-5 (Table 3). The predicted probability of a contact being successful was .545 for F-3 and .426 for F-5.

**Table 3.** Estimated results for model coefficients.

	B	SE	Wald	DF	Sig.	Exp (B)
Game Format (F-3)	0.476	0.087	29.828	1	0.000	1.610
Constant	-0.295	0.068	18.752	1	0.000	0.745
Game Format (F-5)	-0.476	0.087	29.828	1	0.000	0.621
Constant	0.181	0.054	11.086	1	0.001	1.199

The second logistic regression model had a predictive accuracy of 70.8%, a sensitivity of 0%, and a specificity of 100%. The Exp(B) value for Game Format was 1.438 (>1 and significant), indicating that the probability of a contact involving dribbling was 1.438 times higher for F-3 than for F-5 (Table 4). The predicted probability of a contact involving dribbling was .321 for F-3 and .247 for F-5.

**Table 4.** Estimated results for model coefficients.

	B	SE	Wald	DF	Sig.	Exp (B)
Game Format (F-3)	0.364	0.097	13.959	1	0.000	1.438
Constant	-1.112	0.078	202.96	1	0.000	0.329
Game Format (F-5)	-0.364	0.097	13.959	1	0.000	0.695
Constant	-0.749	0.058	166.34	1	0.000	0.473

The third simple logistic regression model had a predictive accuracy of 86.8%, a sensitivity of 0%, and a specificity of 100%. The Exp(B) value for Game Format was 2.553 (>1 and significant), indicating that the probability of a contact involving a shot or continuation of attack with dribbling was 2.553 times higher for F-3 than for F-5 (Table 5). Finally, the predicted probability of a contact involving dribbling plus a shot at goal or continuation of attack was .168 for F-3 and .073 for F-5.

**Table 5.** Estimated results for model coefficients.

	B	SE	Wald	DF	Sig.	Exp (B)
Game Format (F-3)	0.937	0.148	40.128	1	0.000	2.553
Constant	-2.530	0.129	385.37	1	0.000	0.080
Game Format (F-5)	-0.937	0.148	40.128	1	0.000	0.392
Constant	-1.593	0.072	484.89	1	0.000	0.203

## Discussion and conclusions

The high kappa values obtained in our analysis of agreement confirm the reliability of the data used for the current study. The kappa statistic is one of the methods recommended by Blanco-Villaseñor and Anguera (2003) for analyzing intraobserver agreement. The high level of agreement observed is similar to levels reported by other studies that have used the Football Observation System (Camerino, Chaverri, Anguera & Jonsson, 2012; Jonsson et al., 2006; Lapresa et al., 2010) and confirms the reliability of the data used in the analyses required to meet the aims of the study (Reina-Gómez, Hernández-Mendo & Fernández-García, 2010).

We constructed a multiple logistic regression model to investigate the probability of a move being successful (ball reaching the goal area) according to the sector in which the move was initiated and the game format used (F-3 vs F-5) in children aged 6-8 years. In both F-3 and F-5, the probability of the ball reaching the goal area decreased the further away the move was initiated. This observation is consistent with results reported by Castelo (1999), Castellano (2000) and Perea (2008) for soccer and by Ardá (1998) and Arana (2011) for grassroots soccer. The likelihood of success was higher for F-3 than for F-5 in all cases analyzed. Furthermore, the chances of the ball reaching the goal area was 2.426 times higher in F-3 than F-5. A greater use of space (observed in

F-3 in our study) is associated with better quality of play, as indicated by Vales (1998) in a study of tactical indicators of performance in soccer.

We also constructed three simple logistic regression models to analyze the quality of technical actions executed by children aged 6-8 in F-3 and F-5. In the first model, we saw that the likelihood of a contact being effective was 1.610 times higher in F-3 than in F-5. In the second model, we categorized the criterion variable as Adapted when the contact involved dribbling and, as Not Adapted, when it did not. We distinguished between dribbling and non-dribbling, as it has been widely acknowledged that dribbling skills are important in children of a similar age to those in our study (Benedek, 2001; Casal & Ardá, 2003; Lapresa et al., 2005; Lapresa et al., 2008; Vegas 2006; Wein, 1995; Wickstrom, 1990). We found that the smaller game format was better suited to the needs of the children in this age group as players were 1.438 times more likely to dribble the ball in F-3 compared with F-5. In the third model, we also saw that F-3 was more suited to the needs of children as the likelihood of a contact involving both dribbling and a shot at goal or continuation of attack was 2.553 higher in this format.

Our findings show that compared with F-5, F-3 played by children aged 6-8 according to the rules described by Lapresa et al. (2008) is more likely to satisfy competence needs (Ramis, Torregosa, Viladrich & Cruz, 2013) in terms of the use of space (depth of play), effectiveness of play, and development of technical skills. The use of a more suitable game format (F-3 in this case) increases the chances of better play and consequently the chances of children mastering better technical and tactical skills and acquiring more competence (Ardá, 1998; Carvalho & Pacheco 1990; Casal & Ardá 2003; De la Vega et al. 2008; Escudero & Palau 2005; Lapresa, Arana, Anguera & Garzón 2013; Pacheco 2007; Vegas 2006).

We have shown how multiple and simple logistic regression models can be used in observational methodology, and more specifically in studies analyzing how football and sport in general can be adapted to the needs of children using multiple dichotomous variables in addition to those used in our study.

**Note.-** The authors gratefully acknowledge the support of the Spanish government project Observación de la interacción en deporte y actividad física: Avances técnicos y metodológicos en registros automatizados cualitativos-cuantitativos (Secretaría de Estado de Investigación, Desarrollo e Innovación del Ministerio de Economía y Competitividad) during the period 2012–2015 [Grant DEP2012-32124].

## References

- Ábalo, R. (2012). Predicción de lesiones y regresión logística. *Lecturas: Educación Física y Deportes*, 169. Retrieved November 28, 2013 from <http://www.efdeportes.com/efd169/prediccion-de-lesiones-y-regresion-logistica.htm>.
- Álvarez, R. (2007). *Estadística aplicada a las Ciencias de la Salud*. Madrid: Ediciones Díaz de Santos.
- Anguera, M.T. (1979). Observational Typology. *Quality & Quantity. European-American Journal of Methodology*, 13(6), 449-484.
- Anguera, M.T., Arnau, J., Ato, M., Martínez-Arias, R., Pascual, J. & Vallejo, G. (1995). *Métodos de investigación en psicología*. Madrid: Síntesis.

- Anguera, M.T., Blanco-Villaseñor, A. & Losada, J. (2001). Diseños observacionales, cuestión clave en el proceso de la metodología observacional. *Metodología de las Ciencias del Comportamiento*, 3, 135-161.
- Anguera, M.T., Blanco-Villaseñor, A., Hernández-Mendo, A. & Losada, J.L. (2011). Diseños observacionales: ajuste y aplicación en psicología del deporte. *Cuadernos de Psicología del Deporte*, 11(2), 63-76.
- Anguera, M.T., Blanco-Villaseñor, A., Losada, J., Ardá, T., Camerino, O., Castellano, J. & Hernández-Mendo, A. (2003). Instrumento de codificación y registro de la acción de juego en fútbol (SOF-1). *Revista Digital de Alto Rendimiento en Fútbol*, Universidad de Extremadura.
- Anguera, M.T., Magnusson, M.S. & Jonsson, G.K. (2007). Instrumentos no estándar. *Avances en Medicina*, 5(1), 63-82.
- Anguera, M.T. & Hernández-Mendo, A. (2013). La metodología observacional en el ámbito del deporte. *E-balonmano.com: Revista de Ciencias del Deporte*, 9(3), 135-161.
- Araguayo, E., Mena, L.E. & Vavarello, J.R. (2004). Un modelo predictivo para el logro del rendimiento motor. *Lecturas, Educación Física y Deportes*, 10, (73). Retrieved November 6, 2013 from <http://www.efdeportes.com/efd73/rendim.htm>.
- Arana, J. (2011). *Adaptando el fútbol al niño de 12-13 años: análisis observacional de la utilización del espacio en las modalidades de F-7, F-9 y F-11*. Tesis Doctoral: Logroño: Universidad de La Rioja.
- Arana, J., Lapresa, D., Anguera, M.T. & Garzón, B. (2013). Adapting football to the child: an application of the logistic regression model in observational methodology. *Quality & Quantity, International Journal of Methodology*, 47(6), 3473-3480.
- Ardá, A. (1998). *Análisis de patrones en fútbol 7*. Tesis Doctoral. La Coruña: Universidad de La Coruña.
- Ardá, T., Manciro, R., Rial, A., Losada, J.L. & Casal, C. (2014). Análisis de la eficacia de los saques de esquina en la Copa del Mundo de Fútbol 2010. Un intento de identificación de variables explicativas. *Revista de Psicología del Deporte*, 23(1), 165-172.
- Ato, M. & López, J.J. (1996). *Análisis estadístico para datos categóricos*. Madrid: Síntesis.
- Bagley, S.C., White, H. & Golomb, B.A. (2001). Logistic regression in the medical literature: standards for use and reporting, with particular attention to one medical domain. *Journal of Clinical Epidemiology*, 54, 979-985.
- Bakeman, R. & Quera, V. (1995). *Analyzing interaction: Sequential analysis with SDIS and GSEQ*. Cambridge, UK: Cambridge University Press.
- Bakeman, R. & Quera, V. (2001). Using GSEQ with SPSS. *Metodología de las ciencias del Comportamiento*, 3(2), 195-214.
- Bakeman, R. & Quera, V. (2011). *Sequential Analysis and Observational Methods for the Behavioral Sciences*. Cambridge: University Press.
- Benedek, E. (2001). *Fútbol Infantil*. Barcelona: Paidotribo.
- Blanco-Villaseñor, A. & Anguera, M.T. (2003). Calidad de los datos registrados en el ámbito deportivo. En A. Hernández-Mendo (Coord.), *Psicología del Deporte (Vol. 2). Metodología* (p. 35-73). Buenos Aires: Efdportes ([www.efdeportes.com](http://www.efdeportes.com)).
- Broc, M.A. & Gil, C. (2008). Predicción del rendimiento académico en alumnos de ESO y Bachillerato mediante el inventario clínico para adolescentes de Millon (Escala MACI). *Anales de Psicología*, 24(1), 158-167.
- Camerino, O., Chaverri, J., Anguera, M.T. & Jonsson, G.K. (2012). Dynamics of the game in soccer: Detection of T-patterns. *European Journal of Sport Science*, 12(3), 216-224.
- Carvalho, J. & Pacheco, R. (1990). Enseño do futebol: futebol de 11 ou futebol de 7. *Revista Horizonte*, 5(25), 7-11.
- Casal, C.A. (2010). *Análisis observacional de la fase ofensiva del fútbol de alto nivel*. Tesis Doctoral. La Coruña: Universidad de La Coruña.
- Casal, C.A. & Ardá, T. (2003). *Metodología de la enseñanza del fútbol*. Barcelona: Paidotribo.
- Castellano, J. (2000). *Observación y análisis de la acción de juego en el fútbol*. Tesis Doctoral. Vitoria: Universidad del País Vasco.
- Castelo, J. (1999). *Fútbol. Estructura y dinámica del juego*. Barcelona: Inde.
- Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 20, 37-46.
- De Irala, J. (1999). *Regresión logística aplicada a las ciencias de la salud*. Pamplona: Newbook Ediciones.
- De la Vega, R., Del Valle, M.S., Maldonado, A. & Moreno, A. (2008). Una nueva herramienta para la comprensión táctica en fútbol. *Revista Internacional de Medicina y Ciencias de la Actividad Física y del Deporte*, 8, 130-145.
- Escudero, J. M. & Palau, J. M. (2005). Incidencia de la modalidad de juego (fútbol 7 y fútbol 11) sobre la eficacia de las acciones de juego en categorías de formación (11-12 años). *Lecturas: Educación Física y Deportes*, 10 (90). Retrieved March 2, 2014 from <http://www.efdeportes.com/efd90/futbol.htm>.
- Fernández, A., de la Cruz Márquez, J.C., Cueto, B., Salazar, S. & de la Cruz Campos, J.C. (2008). Predicción de lesiones deportivas mediante modelos matemáticos. *Apunts Medicina de l'Esport*, 43(157), 41-44.
- Hellevik, O. (2007). Linear versus logistic regression when the dependent variable is a dichotomy. *Quality & Quantity, International Journal of Methodology*, 43(1), 59-74.
- Herrera, A.N. & Gómez, J. (2008). Influence of equal or unequal comparison group sample sizes on the detection of differential item functioning using the Mantel-Haenszel and logistic regression techniques. *Quality & Quantity, International Journal of Methodology*, 42(6), 739-755.
- Jonsson, G.K. (2006). SOF-coder: Technological and multimedia system for recording data in soccer. En J. Castellano, L. Sautu, A. Blanco-Villaseñor, A. Hernández-Mendo, A. Goñi & F. Martínez de Ilarduya (Coord.). *Socialización y deporte: Revisión Crítica* (pp. 291-300). Vitoria: Diputación Foral de Álava.
- Jonsson, G.K., Anguera, M.T., Blanco-Villaseñor, A., Losada, J., Hernández-Mendo, A., Ardá, T., Camerino, O. & Castellano, J. (2006). Hidden patterns of play interaction in soccer using SOF-CODER. *Behavior Research Methods, Instruments and Computers*, 38(3), 372-381.
- Landis, J.R. & Koch, G.G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33, 159-174.
- Lapresa, D., Álvarez, L., Arana, J., Garzón, B. & Caballero, V. (2013). Observational analysis of the offensive sequences that ended in a shot by the winning team of the 2010 UEFA Futsal Championship. *Journal of Sports Sciences*, 31(15), 1731-1739.
- Lapresa, D., Arana, J. & Carazo, J. (2005). *Pautas para la adecuación de contenidos en el desarrollo psicomotor de prebenjamines y benjamines*. Logroño: Real Federación Española de Fútbol y Universidad de La Rioja.
- Lapresa, D., Arana, J., Anguera, M.T. & Garzón, B. (2013). Comparative analysis of the sequentiality using SDIS-GSEQ and THEME: a concrete example in soccer. *Journal of Sport Sciences*, 31(15), 1687-1695.
- Lapresa, D., Arana, J., Garzón, B., Egüen, R. & Amatria, M. (2008). *Enseñado a jugar el fútbol. Hacia una iniciación coherente*. Logroño: Real Federación Española de Fútbol y Universidad de La Rioja.
- Lapresa, D., Arana, J., Garzón, B., Egüen, R. & Amatria, M. (2010). Adaptando la competición en la iniciación del fútbol: estudio comparativo de las modalidades de F-3 y F-5 en la categoría prebenjamín. *Apunts: Educación Física y Deporte*, 101, 43-56.
- Ortín-Montero, F.J., Moya-Faz, F. & Garcés de los Fayos, E.J. (2013). Sport Psychology: Some new proposals being developed. *Anales de Psicología*, 29(3), 633-636.
- Pacheco, R. (2007). *Fútbol: la enseñanza y el entrenamiento del fútbol 7*. Barcelona: Paidotribo.
- Perea, A.E. (2008). *Análisis de las acciones colectivas en el fútbol de rendimiento*. Tesis Doctoral. Vitoria: Universidad del País Vasco.
- Ramis, Y., Torregosa, M., Viladrich, C. & Cruz, J. (2013). El apoyo a la autonomía generado por entrenadores, compañeros y padres y su efecto sobre la motivación autodeterminada de deportistas de iniciación. *Anales de Psicología*, 29(1), 1695-2294.
- Rando, B., Blanca, M.J. & Frutos, M.A. (2000). Modelo de regresión logística para la intención de hacerse el carnet de donante de órganos y tejidos. *Psicothema*, 12 (2), 464-469.
- Reina-Gómez, A., Hernández-Mendo, A. & Fernández-García, J. C. (201). Multi-facet design for goal scoring in SOCCER-7. *Quality & Quantity, International Journal of Methodology*, 44(5), 1025-1035.
- Rodrigo, M.F., Molina, J.G., García, R. & Pérez, F. (2012). Efectos de la interacción en la predicción del abandono en los estudios de psicología. *Anales de Psicología*, 28(1), 113-119.
- Romero, C. & Vegas, G. (2003). Situación del fútbol 7 en España. En *III Jornadas Internacionales de Fútbol*. Málaga: Real Federación Española de Fútbol.
- Silva, L.C. & Barroso, I.M. (2004). *Regresión Logística*. Salamanca: La Muralla.
- Steyerberg, E.W., Borsboom, G.J., Van Houwelingen, H.C., Eijkemans, M.J. & Habbema, J.D. (2004). Validation and updating of predictive logistic

- regression models: a study on sample size and shrinkage. *Statistics in Medicine*, 23(16), 2567-2586.
- Steyerberg, E.W., Eijkemans, M.J., Van Houwelingen, H.C., Lee, K.L. & Habbema, J.D. (2000). Prognostic models based on literature and individual patient data in logistic regression analysis. *Statistics in Medicine*, 19(2), 141-160.
- Torres, J. & Rivera, E. (1994). *Juegos y deportes alternativos y adaptados en Educación Primaria*. Granada: Rosillo's.
- Vales, A. (1998). *Propuesta de indicadores tácticos para la optimización de los sistemas de juego en fútbol*. Tesis Doctoral. La Coruña: Universidade da Coruña.
- Vegas, G. (2006). *Metodología de la enseñanza basada en la implicación cognitiva del jugador de fútbol base*. Tesis doctoral. Granada: Universidad de Granada.
- Wein, H. (1995). *Fútbol a la medida del niño*. Madrid: CEDIF.
- Wickstrom, R. (1990). *Patrones motores básicos*. Madrid: Alianza Editorial.

(Article received: 03-12-2013; revised: 26-03-2014; accepted: 02-07-2014)