

APLICACIÓN DE DIFERENTES ESTRATEGIAS PARA EL CONTROL DE CALIDAD DEL DATO DE UNA HERRAMIENTA OBSERVACIONAL EN FÚTBOL FORMACIÓN

**Ibon Etxezarra Escudero, Julen Castellano Paulis
y Oidui Usabiaga Arruabarrena
Universidad del País Vasco (UPV/EHU), España**

RESUMEN: El presente trabajo está contextualizado en el ámbito del fútbol de formación y tiene por objeto la estimación de la calidad del dato de una herramienta de observación diseñada *ad hoc* a partir de dos estrategias: coeficientes Kappas y Teoría de la Generalizabilidad. Dicho instrumento (SOFBAS) está configurado por nueve criterios y 73 categorías, contemplándose aspectos del jugador, del equipo y de la interacción. Después de que 14 observadores siguieran un proceso de formación a partir de un protocolo de observación, se realizó un registro para estimar la calidad del dato. A partir de los registros obtenidos se llevaron a cabo la estimación de los coeficientes *Kappa de Fleiss* y *Kappa de Cohen*, así como el estudio de los componentes de varianza aplicados en el estudio de la generalizabilidad. Los resultados obtenidos, valores de asociación por encima de 0.8 y de precisión en la generalización próximo a 1 para la faceta *observadores*, permiten analizar la calidad del dato desde diferentes estrategias, lo que garantiza una mayor precisión de la evaluación. Finalmente se puede concluir con que la herramienta propuesta es fiable para proceder de manera sistematizada a la codificación y registro de la acción de juego en el fútbol de formación.

PALABRAS CLAVE: Fútbol, formación, metodología observacional, calidad del dato.

Manuscrito recibido: 31/01/2013

Manuscrito aceptado: 11/07/2013

Dirección de contacto: Ibon Etxezarra Escudero. C/ Margarita Nelken 5, 1ºB. 01015 Vitoria-Gasteiz, España.
Correo-e.: ibon.etxeazarra@ehu.es

APPLICATION OF DIFFERENT STRATEGIES FOR DATA QUALITY CONTROL OF AN OBSERVATION TOOL FOR YOUNG SOCCER

ABSTRACT: This work is contextualized in the field of young soccer and aims to estimate the data quality of an observation tool which is *ad hoc* designed using two strategies: Kappa coefficients and Generalizability Theory. This instrument (SOFBAS) is formed by nine requirements and 73 categories, contemplating several aspects related to players, team, and interaction. After 14 observers follow a training process using an observation protocol, a record was conducted to estimate the data quality. Taking the obtained records as the starting point, an estimation of the coefficients of the *Cohen* and *Fleiss' Kappa* was carried out, and the study of the components of variance was applied to the study of generalizability. The obtained results, association values above 0.8 and accuracy in the generalization close to 1 for *observer's* facet, allow us to analyze data quality from different strategies, ensuring greater accuracy of the assessment. To conclude, the suggested tool is reliable to proceed in a systematic way to coding and game action including in young soccer.

KEYWORDS: Young soccer, observational methodology, data quality.

APLICAÇÃO DE ESTRATÉGIAS DIFERENTES PARA O CONTROLE DE QUALIDADE DOS DADOS DE UMA FERRAMENTA DE OBSERVAÇÃO NO FUTEBOL DE FORMAÇÃO

RESUMO: O presente trabalho está contextualizado no âmbito do futebol de formação e tem por objectivo a estimativa da qualidade dos dados de uma ferramenta de observação desenhada *ad hoc* a partir de duas estratégias: coeficientes Kappas e Teoria da generalização. Este instrumento (SOFBAS) está configurado por nove critérios e 73 categorias, contemplando-se aspectos do jogador, da equipa e de interação. Depois de 14 observadores seguirem um processo de formação a partir de um protocolo de observação, realizou-se um registo para estimar a qualidade dos dados. A partir dos registos obtidos levou-se a cabo a estimativa dos coeficientes de Kappa de Fleiss y de Cohen, assim como o estudo dos componentes de variação aplicados no estudo de generalização. Os resultados obtidos, valores de associação acima de 0,8 e de precisão da generalização é próximo a 1 para a faceta de observadores, permite analisar a qualidade dos dados a partir de diferentes estratégias, garantindo maior precisão da avaliação. Finalmente, pode-se concluir com concluir que a

ferramenta proposta é fiável para proceder de maneira sistematizada a codificação e o registo da ação de jogo no futebol de formação.

PALAVRAS-CHAVE: Futebol, formação, metodologia de observação, qualidade dos dados.

El número elevado de investigaciones dirigidas al análisis del rendimiento en el ámbito del fútbol profesional contrasta con el escaso número de trabajos referidos al estudio de estos aspectos en la vertiente más formativa de este deporte, es decir, el fútbol base. Teniendo en cuenta que son muchos los jóvenes que se inician cada día en la práctica del fútbol, es preciso disponer de un mayor conocimiento que contribuya a optimizar la intervención en el proceso de enseñanza-aprendizaje de este deporte.

El uso de la metodología observacional (MO) para la observación y análisis de la acción de juego en los deportes se presenta especialmente pródigo (Etxebeste y Castellano, 2005), desde hace casi dos décadas (Hernández-Mendo, 1996) incorporándose nuevas aplicaciones al análisis de la acción motriz, e integrando en sus trabajos parte de las fecundas posibilidades que brindan los diseños observacionales (Anguera, Blanco-Villaseñor, Hernández-Mendo, y Losada, 2011) en deportes como tenis (Garay, 2003), fútbol (Castellano, 2000; Perea, 2008; Robles, 2012), pelota vasca (Usabiaga, 2005), voleibol (Cayero, 2008), baloncesto (Sautu, 2010), escalada (De Benito, 2011), waterpolo (García, 2009) o boxeo (Balmaseda, 2011), por citar algunas. La MO dispone de multitud de posibilidades para acercar al investigador a conocer más sobre contextos naturales, y ofrece mayores y mejores recursos procedimentales en el estudio de la cotidianidad y de las relaciones conducta-ambiente en la que se establecen (Anguera, 1999). El carácter científico de esta metodología, requiere respetar un exquisito rigor en todas las fases de su proceso. En este sentido, se subraya especialmente lo concerniente a posibles sesgos que pueden aparecer en la investigación (Anguera, Blanco-Villaseñor, Losada y Hernández-Mendo, 2000; Hernández-Mendo y Molina, 2002) por ser el tema central de este trabajo. Al igual que se ha propuesto en trabajos anteriores (Hernández-Mendo y Pachuelo, 2012; Hernández-Mendo, Montoro, Reina y Fernández-García, 2012), se necesita realizar una valoración de la calidad del dato, ya que es esencial que el proceso de observación y registro a partir de la herramienta observacional utilizada tenga pocos errores de medida, muestre estabilidad, consistencia y dependencia en las puntuaciones individuales de las características evaluadas (Blanco-Villaseñor, 1997). Teniendo esto presente, el siguiente paso al diseño y configuración de la herramienta de observación es proceder a la comprobación de la fiabilidad en el uso que los observadores hacen de la misma, ya que si no se tiene en cuenta este aspecto, la validez de la investigación puede verse afectada por mediciones poco fiables.

Para medir esta fiabilidad se utilizan diferentes alternativas estadísticas, no existiendo un consenso sobre los coeficientes que se deberían utilizar. La opción estadística a elegir está en estrecha relación con la naturaleza de los datos que en la investigación se manejen (Blanco-Villaseñor, 1993), siendo los de tipo categórico los datos con los que se suele trabajar en MO. Con el propósito de conocer los procedimientos más habituales en el ámbito observacional aplicado al deporte para la estimación de la fiabilidad de la herramienta de observación, se ha llevado a cabo una revisión no exhaustiva de 18 tesis doctorales fundamentadas en la MO. Concretamente, se han analizado los apartados dedicados al estudio de calidad del dato (Anguera, Blanco-Villaseñor y Losada, 2001) propuestos en cada una de ellas.

Tabla 1

Revisión de los estudios de la calidad del dato en Metodología Observacional. Descripción del tipo, Unidades de observación (UO), individuos de las unidades de observación (IUO), momentos (M), n.º de registros (R), participantes, estadísticos y si los indicadores se estimaron para el conjunto de la sesión (S) o para cada uno de los criterios (C) de la herramienta taxonómica.

Autor (año)	Tipo	UO	IUO	M	R	Participantes	Estadísticos	S/C																																																																																								
Castellano (2000)	Inter	2	2	2	4	45' partido de fútbol	P, TK y S, KC, TG (C/O)	S y C																																																																																								
	Intra	2	2	2	4				Garay (2003)	Inter	2	1 y 2	1	2	2 sets de tenis	P, TK y S, KC, TG (C/O)	S y C	Intra	1	1	2	2	Peinado (2004)	Inter	3	ne	1	3	4 periodos de baloncesto*	KC	S	Usabiaga (2005)	Inter	2	2	1	2	2 partidos de pelota**	P, TK y S, KC, TG (MPCS/O)	S y C	Intra	1	2	2	2	Aragundi (2006)	Inter	2	3	2	2	1 partido de voleibol	P, TK y S, KC, TG (C/O)	S y C	Cayero (2008)	Inter	2	2	1	2	4 sets	P, TK y S, KC, TG (C/O)	S y C	Intra	1	2	2	2	Gil (2008)	Inter	4	1	1	4	ne	KC	S	Perea (2008)	Inter	2	2	2	4	45' partido de fútbol	P, TK y S, KC, TG (CZ/O)	S y C	Intra	2	2	2	4	García (2009)	Inter	7	1	1
Garay (2003)	Inter	2	1 y 2	1	2	2 sets de tenis	P, TK y S, KC, TG (C/O)	S y C																																																																																								
	Intra	1	1	2	2				Peinado (2004)	Inter	3	ne	1	3	4 periodos de baloncesto*	KC	S	Usabiaga (2005)	Inter	2	2	1	2	2 partidos de pelota**	P, TK y S, KC, TG (MPCS/O)	S y C	Intra	1	2	2	2	Aragundi (2006)	Inter	2	3	2	2	1 partido de voleibol	P, TK y S, KC, TG (C/O)	S y C	Cayero (2008)	Inter	2	2	1	2	4 sets	P, TK y S, KC, TG (C/O)	S y C	Intra	1	2	2	2	Gil (2008)	Inter	4	1	1	4	ne	KC	S	Perea (2008)	Inter	2	2	2	4	45' partido de fútbol	P, TK y S, KC, TG (CZ/O)	S y C	Intra	2	2	2	4	García (2009)	Inter	7	1	1	7	2 partidos de Waterpolo	Concordancia, % de acuerdo	S y C										
Peinado (2004)	Inter	3	ne	1	3	4 periodos de baloncesto*	KC	S																																																																																								
Usabiaga (2005)	Inter	2	2	1	2	2 partidos de pelota**	P, TK y S, KC, TG (MPCS/O)	S y C																																																																																								
	Intra	1	2	2	2				Aragundi (2006)	Inter	2	3	2	2	1 partido de voleibol	P, TK y S, KC, TG (C/O)	S y C	Cayero (2008)	Inter	2	2	1	2	4 sets	P, TK y S, KC, TG (C/O)	S y C	Intra	1	2	2	2	Gil (2008)	Inter	4	1	1	4	ne	KC	S	Perea (2008)	Inter	2	2	2	4	45' partido de fútbol	P, TK y S, KC, TG (CZ/O)	S y C	Intra	2	2	2	4	García (2009)	Inter	7	1	1	7	2 partidos de Waterpolo	Concordancia, % de acuerdo	S y C																																	
Aragundi (2006)	Inter	2	3	2	2	1 partido de voleibol	P, TK y S, KC, TG (C/O)	S y C																																																																																								
Cayero (2008)	Inter	2	2	1	2	4 sets	P, TK y S, KC, TG (C/O)	S y C																																																																																								
	Intra	1	2	2	2				Gil (2008)	Inter	4	1	1	4	ne	KC	S	Perea (2008)	Inter	2	2	2	4	45' partido de fútbol	P, TK y S, KC, TG (CZ/O)	S y C	Intra	2	2	2	4	García (2009)	Inter	7	1	1	7	2 partidos de Waterpolo	Concordancia, % de acuerdo	S y C																																																								
Gil (2008)	Inter	4	1	1	4	ne	KC	S																																																																																								
Perea (2008)	Inter	2	2	2	4	45' partido de fútbol	P, TK y S, KC, TG (CZ/O)	S y C																																																																																								
	Intra	2	2	2	4				García (2009)	Inter	7	1	1	7	2 partidos de Waterpolo	Concordancia, % de acuerdo	S y C																																																																															
García (2009)	Inter	7	1	1	7	2 partidos de Waterpolo	Concordancia, % de acuerdo	S y C																																																																																								

Autor (año)	Tipo	UO	IUO	M	R	Participantes	Estadísticos	S/C
Landa (2009)	Inter	11	1	1	11	1 set voleibol	KC, KF	S y C
Sautu (2009)	Inter	2	2	2	2	1 partido de baloncesto	P, TK y S, KC, TG (C/O)	S y C
	Intra	1	2	2	2			
Arana (2011)	Inter	3	1	2	3	6 partidos de fútbol	KC, TG	S
Balmaseda (2011)	Inter	2	2	2	4	1 combate de boxeo	KC, TG (SA/O)	S y C
	Intra	2	2	2	4			
De Benito (2011)	Inter	5	1	2	10	1 ascenso en escalada	KC	C
	Intra	5	1	2	10			
Lago (2011)	Inter	4	1	1	2	1a parte partido fútbol	KF	S y C
	Intra	1	1	2	2			
Reina (2011)	Intra	1	1	2	2	2 partidos de fútbol	TG (C/O)	S y C
Montoro (2012)	Inter	2	3	1	2	200 bloqueos de voleibol	P, TK y S, KC, TG (C/O)	S y C
	Intra	1	3	2	2			
Robles (2012)	Inter	Resultados ofrecidos FIFA				13 partidos	P	S y C
	Intra	1	3	2	2	45' partido		

Nota: inter es inter-observadores; intra es intra-observadores; P es Pearson; TK es Tau de Kendall; S es Spearman; KC es Kappa de Cohen; KF es Kappa de Fleiss; CCC es Coeficiente de Concordancia Canónica; TG es Teoría de la Generalizabilidad; (los planes de medida) ne es no específica, *codifica las acciones de dos jugadores en cuatro periodos de un mismo partido; **codifica los 195 primeros golpes en dos partidos diferentes.

El estadístico más empleado por aportar un índice de asociación es el coeficiente *Kappa de Cohen* (Bakeman y Gottman, 1989), empleado en 16 de las 18 tesis analizadas (Aragundi, 2006; Arana, 2011; Balmaseda, 2011; Castellano, 2000; Cayero, 2008; De Benito, 2011; Garay, 2003; Gil, 2009; Lago, 2011; Landa, 2008; Montoro, 2012; Peinado, 2004; Perea, 2008; Robles, 2012; Sautu, 2009; Usabiaga, 2005). Se trata de un estadístico que permite la comparación de los registros punto por punto, respetando el orden de registro de las conductas. De este modo tiene en cuenta y corrige el efecto del azar. Por otro lado, en cinco de los trabajos de tesis revisados se plantean análisis de correlación como complemento del coeficiente *Kappa de Cohen* (Aragundi, 2006; Castellano, 2000; Cayero, 2008; Garay, 2003; Montoro, 2012; Perea, 2008; Robles,

2012; Sautu, 2009; Usabiaga, 2005). Para ello se hace uso de los coeficientes de *Pearson*, *Tau de Kendall* y *Spearman*, indicadores que informan sobre la concordancia global entre dos registros, pero el problema que acarrear es que sobrestiman la concordancia debido a que, utilizan las frecuencias totales, además de no considerar el efecto del azar. Mención especial requieren los trabajos de tesis de Landa (2009) y Lago (2011), en los que se plantea la necesidad de establecer el nivel de concordancia global entre más de dos observadores. Para determinarlo proponen aplicar el coeficiente *Kappa de Fleiss* (Fleiss, 1971), estadístico que permite calcular el nivel de concordancia global entre tres o más unidades de observación.

Concluimos la presente revisión destacando el número considerable de tesis que utilizan la teoría de la generalizabilidad como complemento a los anteriores análisis estadísticos (Aragundi, 2006; Arana, 2011; Balmaseda, 2011; Castellano, 2000; Cayero, 2008; Garay, 2003; Montoro, 2012; Perea, 2008; Reina, 2011; Robles, 2012; Sautu, 2009; Usabiaga, 2005), con el fin de integrar diferentes fuentes de variación tales como observadores distintos, diversas ocasiones, varios instrumentos, tipos variados de registro, ocasiones diversas, etc., en una estructura global (Blanco-Villaseñor, 1997).

A partir de lo expuesto anteriormente, el presente trabajo tiene por objeto la estimación desde diferentes estrategias del análisis de la calidad del dato de una herramienta de observación diseñada *ad hoc* aplicada en el fútbol formativo. Tras comprobar su fiabilidad va a ser posible implementarla con el objetivo de conocer el perfil de juego que caracteriza el desempeño técnico-táctico de los jóvenes deportistas en sus diferentes etapas y proponer pautas para la optimización del contexto deportivo.

MÉTODO

Participantes

14 observadores participaron en este estudio, que fueron agrupados por pares. Previamente, los observadores fueron formados en el manejo de la herramienta de observación, y posteriormente, llevaron a cabo la codificación de una parte de un partido de fútbol de categoría infantil. Cada observación supuso aproximadamente un centenar de posesiones individuales que, por cada uno de los nueve criterios que conforman la herramienta de observación, supusieron el registro de alrededor de 1000 códigos para cada una de las observaciones. A partir de los siete registros resultantes se obtuvieron los indicadores de concordancia inter-observadores.

Instrumentos

Herramienta de observación: SOFBAS. Siguiendo la línea de estudios previos (Castellano, Hernández-Mendo, Morales-Sánchez y Anguera, 2007; Hernández-Mendo, Díaz-Martínez y Morales, 2010; Refoyo, Romaris y Sampedro, 2009; Robles, 2012) se procedió a diseñar y optimizar una herramienta expresamente diseñada para la observación

del fútbol de formación (Castellano y Echeazarra, 2013), que sin embargo, únicamente presentaron los valores de la concordancia intraobservadores (con el coeficiente de la Kappa de Cohen), que se situó por encima de 0.90 para cada criterio y de 0.95 para el conjunto de la sesión. Esta herramienta permite analizar aspectos novedosos tales como la conducta del jugador que interviene con el móvil, la superficie de contacto con la que desarrolla la acción con balón, el lugar de inicio y fin de la misma, así como su ubicación en relación al resto de jugadores, entre otros (tabla 2). En la presente investigación se utilizan *datos secuenciales de multievento* (Bakeman y Quera, 1996). De esta manera, se ha registrado una categoría de cada uno de los nueve criterios en cada paso de secuencia de manera concurrente (Anguera et al. 2000), no siendo consideradas las duraciones de las posesiones individuales.

Tabla 2
Criterios configuradores de la herramienta de observación SOFBAS.

Criterio	Descripción y categorías
Inicio	<i>Inicio de la jugada</i> (10 niveles) hace referencia al modo en que el jugador inicia la posesión individual. Se establecieron las categorías: continuar la posesión, recuperar la posesión, interceptar el balón, tomar rechace, sacar de banda, de esquina, de puerta, de falta, lanzar un penalti y bote neutral.
Equipo	<i>Equipo que tiene la posesión del balón</i> (2 niveles). Indica el equipo al que pertenece el jugador que está en posesión del balón. Se establecieron las categorías: equipo 1 y 2.
EJE	<i>Espacio de juego efectivo</i> (4 niveles). Contextualiza la acción del jugador que está en posesión del balón en relación al resto de jugadores, compañeros y contrarios, determinando si se encuentra fuera o dentro del perímetro formado por los jugadores del mismo equipo y del equipo contrario más alejados del centro. Se establecen las categorías: fuera-fuera, fuera-dentro, dentro-dentro, dentro-fuera.
ZonaI	La zona de inicio o <i>ZonaI</i> (12 niveles) es aquella en la que el jugador entra en contacto por primera vez con el balón y la de finalización o <i>ZonaF</i> (12 niveles) aquella en la que lo toca por última vez. El espacio de juego se ha dividido longitudinal y transversalmente, resultando 12 zonas diferentes.
ZonaF	
Cuerpo	El criterio <i>cuerpo</i> determina la parte del cuerpo (5 niveles) con la que el jugador entra en contacto con el balón a lo largo de la posesión individual. Se establecieron las categorías: pierna, tronco, cabeza, mano o mixto (cuando existió una combinación de las anteriores).
Contactos	<i>Contactos con el balón</i> representa el número de contactos (4 niveles) que realiza el jugador en cada una de las posesiones individuales. Las categorías son: 1, 2, 3 y 4 o más contactos.

Criterio	Descripción y categorías
Conducta	Recoge la <i>conducta</i> (10 niveles) desplegada por el jugador con la que da por finalizada su posesión individual. La categorías establecidas son: pase, golpeo, pérdida con continuidad, fuera, falta a favor, falta en contra, tiro, gol, robo y conjunto vacío.
Dirección	<i>Dirección del balón</i> define la dirección (4 niveles) que toma el balón cuando el jugador se desprende de él en relación a los ejes trasversales (izquierda y derecha) y longitudinales (adelante y atrás). La orientación la determina la portería contraria. Las categorías establecidas fueron: izq-ade, izq-atrás, der-ade y der-atrás.

Material

Para el registro de los datos se utilizó la aplicación informática MOTS (Castellano, Perea, Alday y Hernández-Mendo, 2008). Para al análisis estadístico se utilizó el paquete estadístico SPSS para Windows versión 20.0. También se empleó el programa SDIS-GSEQ (Bakeman y Quera, 1996) en versión para Windows para realizar el análisis secuencial, y los programas estadísticos SAS (Schlotzhauer y Littel, 1997; SAS Institute Inc, 1999) y SAGT v1.0 versión 211 (Ramos, Hernández-Mendo, Pastrana y Blanco-Villaseñor, 2012) para el análisis de la generalizabilidad.

Procedimiento

Antes del registro, los observadores recibieron una formación a lo largo de tres semanas que supuso un total de 15 horas de entrenamiento, distribuidas en tres sesiones de hora y media cada una, para lo que se siguió un protocolo de observación expresamente diseñado para tal efecto que incluyó: cronograma de la fase de formación y fecha para la prueba definitiva, esquema básico de la herramienta taxonómica así como videos con ejemplos y contraejemplos de cada una de las categorías del conjunto de los nueve criterios. Los 14 observadores se organizaron en siete pares para garantizar de este modo la concordancia consensuada (Anguera, 1990). Tras ello, y por parejas, procedieron al registro de una parte del partido seleccionado.

Tal y como está configurada la herramienta de observación es probable que pueda ocurrir algún error de omisión en el registro de alguna categoría que pudiese provocar un desajuste en el conjunto de los registros comparados. Esto provocaría una subestimación del índice de concordancia entre los observadores cuando realmente no ha sido tal (Castellano, 2000). Para evitar esto, y antes de proceder al cálculo de los estadísticos se procedió a hacer un filtro, emparejando las acciones de 'balón parado' como son las interrupciones reglamentarias, entre los siete registros, consideradas fáciles de identificar por los observadores. Finalmente, a partir de los siete registros obtenidos se aplicaron los estadísticos pertinentes para estimar la calidad del dato a partir del análisis de la concordancia inter-observadores.

Tratamiento de datos

El cálculo del coeficiente *Kappa* fue implementado para cada uno de los criterios que configuran la herramienta de observación así como para la sesión completa entre los siete pares de observadores. Primero, se compararon de manera global a partir del coeficiente *Kappa de Fleiss* (Fleiss, 1971) válido para medir la fiabilidad de escalas cuantitativas (Fleiss y Cohen, 1973). Los valores de interpretación de este estadístico oscilan entre -1 y +1, siendo $K \leq 0.20$ *pobre*; $K = 0.21-0.40$ *débil*; $K = 0.41-0.60$ *moderada*; $K = 0.61-0.80$ *buena*; $K = 0.81-1.00$ *muy buena* (Fleiss, 1971). En segundo lugar, se compararon los registros de dos en dos, para lo que se utilizó el coeficiente *Kappa de Cohen*. Finalmente, y como complemento de lo anterior, se realizaron los análisis de componentes de varianza y de generalizabilidad para lo que se configuró un diseño de dos facetas, *observador*categorias*.

RESULTADOS

En relación al coeficiente *Kappa de Fleiss*, se estimaron valores de entre 0.63 y 0.85 para el conjunto de los siete grupos de observadores en cada uno de los criterios de la herramienta de observación (*Equipo*: 0.79; *Inicio*: 0.85; *EJE*: 0.64; *Zona I*: 0.70; *Zona F*: 0.63; *Cuerpo*: 0.74; *Contactos*: 0.80; *Conducta*: 0.76; y, *Orientación*: 0.70), lo que supone un nivel de concordancia bueno. Fueron los criterios *Contactos* e *Inicio* los que mayor nivel de acuerdo presentaron, situándose en el máximo rango (0.81-1.00), considerada como muy buena concordancia global, si tenemos en cuenta que los valores de interpretación del coeficiente *Kappa* oscilan entre -1 y +1.

Como complemento al análisis anterior, se estimaron los niveles de concordancia entre cada par de observadores, para identificar un posible déficit de formación en algún grupo de observadores en el uso de SOFBAS. Para ello, se compararon los registros dos a dos (tabla 3) a partir del coeficiente *Kappa de Cohen* (Cohen, 1960), que sigue los mismos intervalos que la *Kappa de Fleiss*. Los valores obtenidos se situaron en el rango 0.68-0.98, lo que constata que el nivel de concordancia entre las distintas parejas de observadores varió sensiblemente.

Tabla 3

Coefficientes Kappa de Cohen interobservadores para el conjunto de los siete pares de observadores.

Inter-observadores						
G1	G2	G3	G4	G5	G6	G7
G1	0.73	0.93	0.91	0.91	0.83	0.74
G2		0.79	0.77	0.77	0.73	0.68
G3			0.85	0.98	0.76	0.70
G4				0.87	0.79	0.73
G5					0.74	0.68
G6						0.69

Nota: G1, par de observadores 1; G2, par de observadores 2; G3, par de observadores 3; G4, par de observadores 4; G5, par de observadores 5; G6, par de observadores 6.

Para finalizar, se procedió a la realización del estudio de generalización (G) a partir de un modelo de dos facetas: *categorías* [O] y *observadores* [C]. El plan de medida fue C/O, es decir, se colocó como faceta diferenciación [C] y en la faceta instrumentación [O]. Se estimó el porcentaje de variabilidad de cada una de las facetas y sus interacciones, así como los coeficientes absolutos y relativos de generalizabilidad (tabla 4).

Tabla 4

*Análisis de los componentes de variancia y de los coeficientes de generalización en el diseño de dos facetas O*C [Observadores * Categorías].*

Observadores* Categorías				
Facetas	g° de l	Pr > F para el modelo <.0001		
		SC (Tipo III)	Pr > F de variancia	%
Observador [O]	6	68.6	1.000	0.1
Categoría [C]	56	50537.4	<.0001	98.2
O*C	336	737.4	1.000	1.7
C / O		$\xi\rho^2_{(\delta)} = 0.81$ y $\xi\rho^2_{(\Delta)} = 0.07$		

Nota. En la tabla figuran el coeficiente de determinación (r^2), los grados de libertad (g° de l), la suma de cuadrados para los datos tipo III (SC tipo III), el % de variabilidad de cada una de las facetas y de sus interacciones así como los coeficientes absolutos y relativos de generalizabilidad ($\xi\rho^2_{(\delta)}$ y $\xi\rho^2_{(\Delta)}$) para el diseño C/O.

La determinación de las fuentes de varianza reveló que casi toda la variabilidad (98 %) estuvo asociada a las facetas *categorías* [C], presentando una prácticamente nula variabilidad la faceta *observadores* [O]. Esto muestra una situación ideal que atiende a que el registro realizado por los siete pares de observadores no se vio afectado, no existiendo diferencias entre los registros. El análisis global de los coeficientes de generalizabilidad reveló que la precisión de generalización de los resultados fue óptima (0.81) cuando la faceta [O] fue ubicada en la faceta de instrumentación.

DISCUSIÓN

El objetivo del estudio fue utilizar diferentes estrategias para estimar la calidad del dato de una herramienta de observación elaborada *ad hoc* y comprobar así su nivel de fiabilidad. El uso de estrategias que engloben al conjunto de los observadores o de dos en dos, pueden dar diferentes perspectivas en el control de la calidad del dato aplicada a la concordancia entre observaciones.

Primeramente destacamos que en el presente trabajo ha sido novedoso el estudio de la fiabilidad a partir del coeficiente *Kappa de Fleiss* como indicador de concordancia global para siete pares de observadores aplicado en el ámbito del fútbol de formación. Los valores obtenidos en el coeficiente *Kappa de Fleiss* superan el 0.70 en casi todos los criterios, quedándose próximos a 0.90. Estos valores se consideran óptimos, siendo similares a los estimados en otros estudios previos (Lago, 2011; Landa, 2009; Usabiaga, Castellano, Blanco-Villaseñor y Casamichana, 2013).

En relación al coeficiente *Kappa de Cohen* para cada par de grupos de observadores, la mayoría mostraron valores aceptables, por encima de 0.75. Estos valores fueron similares a los encontrados en estudios realizados en otros deportes y ámbitos de aplicación (ver tabla 1) y algo superiores a los estimados por Bloomfield, Polman y O'Donoghue (2007) cuando, utilizando el programa *Observer system Version 5.1* (Noldus Information Technology, The Netherlands), codificaron el desplazamiento de jugadores de fútbol de la Premier Ligue (primera división inglesa).

Por último, los valores estimados para la variabilidad aportada por cada faceta (en %) en el análisis de los componentes de variancia así como el coeficiente relativo y absoluto de generalizabilidad, se ajustan a lo requerido en esta fase de la investigación. Por un lado, cabe subrayar que la faceta *observador* asumió muy poca variabilidad en el modelo (próximo a 0), siendo la faceta *categoría* la que aportó casi la totalidad de la misma. Estos valores son similares a los encontrados en las tesis analizadas con anterioridad (p.e. Aragundi, 2006; Cayero, 2008; Sautu, 2009, entre otros). Con esto, se interpreta la no existencia de variabilidad entre los observadores. Por otro lado, con relación a la estimación de la precisión de generalización, el coeficiente de generalizabilidad se aproxima a la unidad para el diseño *C/O* cuando es O la faceta establecida como faceta de instrumentación. Se entiende así que los índices de fiabilidad

inter-observadores son los adecuados para llevar a cabo una investigación a partir de esta herramienta de observación y con estos observadores.

Algunas de las limitaciones que hemos encontrado tienen que ver con la necesidad de implementar un mayor periodo de formación para los observadores, en general, y de manera particular para mejorar el acuerdo en la codificación de los criterios *Zona de finalización* y, en especial, al *Espacio de juego efectivo* (EJE). Este criterio consideramos de especial relevancia debido a su carácter novedoso, ya que permite contextualizar la acción del jugador con posesión de balón en relación al resto de jugadores, compañeros y contrarios. Ante esta circunstancia, y si se optase por comenzar la fase activa de la investigación en el ámbito observacional (Anguera et al., 2000), codificando y analizando, deberíamos ser cautelosos en la interpretación de los resultados de la misma. En este sentido, y previo al desarrollo del registro de las sesiones estipuladas, deberíamos valorar opciones dirigidas a mejorar estos resultados de la fiabilidad (Anguera et al., 2000; Hernández-Mendo y Molina, 2002) entre los que se podrían incluir: la redefinición de las categorías, mejorar y ampliar la calidad y variabilidad de los videos que se utilizaron en la fase de formación de los observadores, intensificar el proceso de formación de los observadores o buscar nuevas alternativas, más fácilmente identificables por los observadores, a la hora de visualizar los aspectos del juego que interese al investigador respecto al uso estratégico del espacio por los equipos.

CONCLUSIÓN

En definitiva, como conclusión podemos afirmar que la utilización de diferentes estrategias para llevar a cabo la calidad de los datos satisface la necesidad del investigador para asegurar que el registro no lleve implícito errores propios. Paralelamente podemos afirmar que SOFBAS se trata de un sistema de observación sencillo, fiable, y diseñado expresamente para analizar y describir el fútbol en las etapas de formación. Sin embargo, debemos ser cautelosos con el uso de algunos criterios que, al presentar valores por debajo de los deseados, deben ser mejorados para poder continuar con las sucesivas fases de la investigación sistematizada (Anguera et al., 2000) o ser utilizados en posteriores estudios. Finalmente hay que destacar que la aplicación de esta herramienta al estudio del fútbol de formación puede abrir nuevas vías con las que optimizar el proceso formativo en este ámbito. Aun así, más estudios en este ámbito de intervención son necesarios para obtener cada vez más y mejor información con las que poder seguir avanzando.

REFERENCIAS

- Anguera, M. T. (1990). Metodología observacional. En J. Arnau, M. T. Anguera y J. Gómez-Benito (Ed.), *Metodología de la investigación en ciencias del comportamiento* (pp. 125-236). Murcia: Universidad de Murcia.
- Anguera, M. T. (1999, 23 de noviembre). *Hacia una evaluación de la actividad y su contexto: ¿Presente o futuro para la metodología?* Discurso de ingreso a la Real Academia de Doctores. Barcelona.
- Anguera, M. T., Blanco-Villaseñor, A., Hernández-Mendo, A., y 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. y Losada, J. L. (2001). Diseños observacionales, cuestión clave en el proceso de la Metodología Observacional. *Metodología de las Ciencias del Comportamiento*, 3(2), 135-160.
- Anguera, M. T., Blanco-Villaseñor, A., Losada, J. L. y Hernández-Mendo, A. (2000). La metodología observacional en el deporte: conceptos básicos. *Lecturas: EF y Deportes. Revista Digital*, 24. Recuperado el 17 de enero de 2001 desde <http://www.efdeportes.com>.
- Aragundi, C. A. (2006). *Observación y análisis de la Colocación en el Voleibol*. Tesis Doctoral sin publicar, Universidade da Coruña, A Coruña, España.
- Arana, J. (2011). *Adaptando el fútbol al niño de 12-13 años: análisis observacional de la utilización del espacio en las modalidades de fútbol 7, fútbol 9 y fútbol 11*. Tesis Doctoral sin publicar, Universidad de La Rioja, Logroño, España.
- Bakeman, R. y Gottman, J. M. (1989). *Observación de la interacción: introducción al análisis secuencial*. Madrid: Morata.
- Bakeman, R. y Quera, V. (1996). *Análisis de la interacción*. Madrid: RAMA.
- Balmaseda, M. (2011). *Análisis de las acciones técnico-tácticas del boxeo de rendimiento*. Tesis Doctoral sin publicar, Universidad del País Vasco, Vitoria, España.
- Blanco-Villaseñor, A. (1993). Fiabilidad, precisión, validez y generalización de los diseños observacionales. En M. T. Anguera (Ed.), *Metodología Observacional en la investigación psicológica* (Vol. II, pp.149-274). Barcelona: PPU.
- Blanco-Villaseñor, A. (1997). *Metodologies qualitatives en la investigació psicològica*. Barcelona: Edicions de la Universitat Oberta de Catalunya.
- Bloomfield, J., Polman, R. y O'Donoghue, P. (2007). Physical demands of different positions in FA Premier League soccer. *Journal of Sports Science and Medicine*, 6(1), 63-70.
- Castellano, J. (2000). *Observación y análisis de la acción de juego en el fútbol*. Tesis Doctoral sin publicar, Universidad del País Vasco, San Sebastián. España.

- Castellano, J., Hernández-Mendo, A., Morales-Sánchez, V. y Anguera, M. T. (2007). Optimising a probabilistic model of the development of play in soccer. *Quality & Quantity*, 41(1), 93-104.
- Castellano, J., Perea, A., Alday, L. y Hernández-Mendo, A. (2008). Measuring and observation tool in sports. *Behavior Research Methods*, 40, 898-903.
- Castellano, J. y Etxezarra, I. (2013). Adapting model competition in youth football: a comparative study of 5-a side football and 7-a side football in U-9 players. En H. Nunome, B. Drust y D. Dawson (Ed.), *Science and football VII* (pp. 311-316). London: Routledge.
- Cayero, R. (2008). *Observación de la acción de juego del Voleibol: Análisis secuencial y de variabilidad*. Tesis doctoral sin publicar, Universidad de Málaga, Málaga, España.
- Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 41, 687-699.
- De Benito, A. M. (2011). *Análisis y cuantificación de las acciones técnicas en la escalada deportiva de alto nivel y sus implicaciones musculares*. Tesis doctoral sin publicar, Universidad de Valladolid, Valladolid, España.
- Etxebeste, J., y Castellano, J. (2005). Avances en los estudios sobre la acción motriz. En Á. M. González (Ed.), *Avances en Ciencias del Deporte* (pp. 23-43). Bilbao: Universidad del País Vasco.
- Fleiss, J. (1971). Measuring nominal scale agreement among many raters. *Psychological Bulletin*, 76(5), 378-382.
- Fleiss, J. L. y Cohen, J. (1973). The equivalence of weighted Kappa and the intraclass correlation coefficient as measures of reliability. *Educational and Psychological Measurement*, 33, 613-619.
- Garay, J. O. (2003). *Observación y análisis de la acción de juego del tenis de dobles*. Tesis Doctoral sin publicar, Universidad de Málaga, Málaga, España.
- García, P. (2009). *Evaluación cuantitativa de la desigualdad numérica temporal simple con posesión mediante observación sistemática en water polo*. Tesis doctoral sin publicar, Universidad Autónoma de Madrid, Madrid, España.
- Gil, A. (2008). *Los porteros de fútbol, ¿se comportan como sistemas complejos? Estudio comparativo de Iker Casillas y Víctor Valdés*. Tesis Doctoral sin publicar, Universitat de Barcelona, Barcelona, España.
- Hernández-Mendo, A. (1996). *Observación y análisis de patrones de juego en deportes socio-motrices*. Tesis Doctoral sin publicar, Universidad de Santiago de Compostela, Santiago de Compostela, España.
- Hernández-Mendo, A. y Molina, M. (2002). Cómo usar la observación en la psicología del deporte: principios metodológicos. *Lecturas: EF y Deportes*, 49. Recuperado desde <http://www.efdeportes.com/efd49/obs.htm>

- Hernández-Mendo, A. y Pachuelo, L. (2012). Una herramienta observacional para la evaluación del desarrollo moral en las clases de educación física en primaria. *Revista Iberoamericana de Psicología del Ejercicio y el Deporte*, 7(2), 287-306.
- Hernández-Mendo, A., Díaz-Martínez, F. y Morales, V. (2010). Construcción de una herramienta observacional para evaluar las conductas prosociales en las clases de educación física. *Revista de Psicología del Deporte*, 19(2), 305-318.
- Hernández-Mendo, A., Montoro, J., Reina, A. y Fernández-García, J. C. (2012). Desarrollo y optimización de una herramienta observacional para el bloqueo en voleibol. *Revista Iberoamericana de Psicología del Ejercicio y el Deporte*, 7(1), 15-31.
- Lago, J. (2011). *Influencia de los condicionantes estratégicos y las variables situacionales en el rendimiento de la fase ofensiva en fútbol*. Tesis Doctoral sin publicar, Universidad de Vigo, Pontevedra, España.
- Landa, R. (2009). *Análisis de las categorías y de la fiabilidad interobservadores en los sistemas de observación que evalúan el rendimiento de las acciones de juego en voleibol, dentro de la división de honor femenina española*. Tesis Doctoral. Granada. Universidad de Granada.
- Montoro, J. (2012). *Observación y análisis del bloque en voleibol*. Tesis Doctoral sin publicar, Universidad de Málaga, Málaga, España.
- Peinado, L. (2004). *Processos cognoscitius que intervenen en la millora de la comunicació motriu*. Tesis Doctoral. Barcelona. Universitat de Barcelona.
- Perea A. (2008). *Análisis de las acciones colectivas en el fútbol de rendimiento*. Tesis Doctoral sin publicar, Universidad del País Vasco, San Sebastián, España.
- Ramos, F. J., Hernández-Mendo, A., Pastrana, J. L. y Blanco-Villaseñor, A. (2012). *SAGT: Software para la Aplicación de la Teoría de la Generalizabilidad*. Proyecto fin de carrera para la titulación: Ingeniería Técnica en Informática de Gestión de la Escuela Técnica Superior de Ingeniería Informática, Universidad de Málaga, España.
- Refoyo, I., Romaris, I. U. y Sampedro, J. (2009). Analysis of men's and women's basketball fast-breaks. *Revista de Psicología del Deporte*, 18(3), 439-444.
- Reina, A. (2011). *Evaluación de factores implicados en el rendimiento del fútbol profesional*. Tesis Doctoral sin publicar, Universidad de Málaga, Málaga, España.
- Robles, F. J. (2012). *Observación y análisis de las acciones ofensivas de la selección española de fútbol en la Eurocopa de 2008 y en el Mundial de 2010*. Tesis Doctoral sin publicar, Universidad del País Vasco, Vitoria, España.
- Robles, F. J. y Castellano, J. (2012). Comparación entre el juego ofensivo de la selección española de fútbol y sus rivales en la Eurocopa'08 y Mundial'10. *Revista Iberoamericana de Psicología del Ejercicio y el Deporte*, 7(2), 323-338.
- SAS Institute Inc. (1999). *SAS/STAT User's Guide*. Version 7-1. Cary, NC: SAS Institute Inc.

- Sautu, L. (2009). *Observación y Análisis de la acción del juego en baloncesto ACB*. Tesis Doctoral sin publicar, Universidad del País Vasco, Vitoria, España.
- Schlotzhauer, S. D. y Littlell, R. C. (1997). *SAS System for Elementary Statistical Analysis*, Second Edition, Cary, NC: SAS Institute Inc.
- Usabiaga, O. (2005). *Descripción y análisis de la pelota vasca: aplicación en mano parejas*. Tesis Doctoral sin publicar, Universidad del País Vasco, San Sebastián, España.
- Usabiaga, O., Castellano, J., Blanco-Villaseñor, A. y Casamichana, D. (2013). La Teoría de la Generalizabilidad en las primeras fases del método observacional aplicado en el ámbito de la iniciación deportiva: calidad del dato y estimación de la muestra. *Revista de Psicología del Deporte*, 22(1), 103-109.