

# Estudio de fiabilidad de los indicadores de rendimiento en la liga asobal

## Reliability study of performance indicators in asobal league

### Estudo da fidelidade dos indicadores de rendimento na Liga Asobal

Andrea Blanco<sup>1</sup>, Sergio José Ibáñez<sup>1\*</sup>, Antonio Antúñez<sup>2</sup>, Antonio Hernández-Mendo<sup>3</sup>

<sup>1</sup>Facultad de Ciencias del Deporte. Universidad de Extremadura, <sup>2</sup>Facultad de Educación. Universidad de Extremadura,

<sup>3</sup>Facultad de Ciencias de la Educación. Universidad de Málaga

**Resumen:** El objetivo de esta investigación fue validar los indicadores de rendimiento ofrecidos por las denominadas estadísticas oficiales de juego de la liga ASOBAL. Se realizó un muestreo de eventos sobre tres partidos ( $n=7950$ ). Las variables utilizadas fueron los criterios que conforman el sistema de observación (localización, zona de lanzamiento, forma de finalización del lanzamiento, sanciones disciplinarias, faltas, acciones de ataque, acciones de defensa, zonas de la portería y eficacia del portero), el observador oficial encargado de elaborar las estadísticas, y dos observadores entrenados para el estudio. La recogida de datos se realizó mediante un sistema de formatos de campo, utilizado por los dos observadores entrenados. El análisis de calidad del dato confirma la fiabilidad (objetividad) de la muestra analizada. El estudio de generalizabilidad manifiesta la precisión de medida ( $\phi=1$ ;  $\rho^2=.999$ ), el adecuado funcionamiento del sistema de observación ( $\phi=.000$ ;  $\rho^2=.000$ ) y la capacidad de generalización ( $\phi=.998$ ;  $\rho^2=.998$ ). Finalmente el estudio de decisión donde se proponen tres nuevos diseños con sobre la base de los datos del diseño original (variación del número de observadores, de criterios y partidos), certificó que los resultados obtenidos en el estudio de generalizabilidad del diseño original no eran mejorables con los diseños alternativos. Se concluye que los indicadores de rendimiento analizados son fiables, pudiendo generalizar los resultados de forma precisa para la realización de trabajos de investigación.

**Palabras Clave:** Metodología observacional, fiabilidad, Teoría de la Generalizabilidad, balonmano.

**Abstract:** The aim of this research was to validate the performance indicators offered by the so-called game official statistics of the ASOBAL league. A sampling of events in three matches ( $n=7950$ ). The variables used were the criteria which make up the observation system (placing, throwing area, way of finishing throwing, disciplinary measures, fouls, offensive actions, defense actions, goal areas and goalkeeper efficiency), the official observer to compile statistics and two trained observers for the study. The data capture was performed using a system of field formats used by two trained observers. The agreement study proves the (objectivity) reliability of the analyzed sampling. The Generaliza-

bility study shows the reliable accuracy of it ( $\phi=1$ ;  $\rho^2=.999$ ), the appropriate functioning of the observation system ( $\phi=.000$ ;  $\rho^2=.000$ ) and the accuracy of generalization of the results ( $\phi=.998$ ;  $\rho^2=.998$ ). Finally, the study of decision, where three new designs based on the original design data are proposed (variation of the number of observers, criteria and matches), certified that the results obtained in the study generalizability of the original design, not were improvable with alternative designs. To sum up, the performance indicators are reliable and the results can be used accurately for research works.

**Key words:** observational methodology, reliability, Theory of Generalizability, handball.

**Resumo:** O objetivo desta investigação foi validar os indicadores de rendimento oferecidos pelas denominadas estatísticas oficiais de jogo da Liga ASOBAL. Foi realizada uma amostragem de eventos em três jogos ( $n=7950$ ). As variáveis utilizadas foram os critérios que englobam o sistema de observação (localização, zona de lançamento, forma de finalização do lançamento, sanções disciplinares, faltas, acções de ataque, acções de defesa, zonas de baliza e eficácia do guarda-redes), o observador oficial encarregue de elaborar as estatísticas, e dois observadores treinados para o estudo. A recolha de dados foi realizada mediante um sistema de formatos de campo, utilizado pelos dois observadores. A análise da qualidade de dados confirma a fidelidade (objectividade) da amostra analisada. O estudo da generalização manifesta precisão de medida ( $\phi=1$ ;  $\rho^2=.999$ ), o adequado funcionamento do sistema de observação ( $\phi=.000$ ;  $\rho^2=.000$ ) e a capacidade de generalização ( $\phi=.998$ ;  $\rho^2=.998$ ). Por último, o estudo de decisão a partir do qual se propõem três novos delineamentos tendo por base o delineamento original (variação do número de observadores, de critérios e jogos), certificou que os resultados obtidos no estudo de generalização do delineamento original não eram melhoráveis com delineamentos alternativos. Conclui-se que os indicadores de rendimento analisados são fiáveis, podendo generalizar os resultados de forma precisa para a realização de trabalhos de investigação.

**Palavras-chave:** Metodologia observacional, Fidelidade, Teoria da Generalização, Andebol.

## Introducción

Actualmente, en los deportes colectivos de máxima práctica y difusión mediática (p.e. fútbol, baloncesto o balonmano), se suministran indicadores fruto de la observación, definidos en el ámbito científico como indicadores de rendimiento (Hughes y Bartlett, 2002, Prieto, 2011). Su importancia en los deportes

colectivos va en aumento porque ayuda a la toma de decisiones directivas por parte de los entrenadores, ofreciendo una aproximación más objetiva a la realidad del juego tanto en competición amateur (García, Ibáñez, Parejo, Feu, y Cañadas, 2011), como profesional (Ohnjec, Vuleta, Milanović, y Gruić, 2008; Volossovitch, Dumangane, y Rosati, 2012).

Los datos que se emplean en las investigaciones para el análisis del juego provienen habitualmente de las instituciones que organizan las competiciones (Sampaio, Ibáñez y Lorenzo,

Dirección para correspondencia [Correspondence address]: Sergio José Ibáñez Godoy. Facultad de Ciencias del Deporte. Universidad de Extremadura. Avd. de la Universidad s/n. 10071 Cáceres (España). E-mail: [sibanez@unex.es](mailto:sibanez@unex.es)

2013). Para ello, seleccionen un grupo de analistas que son los encargados de su registro. Estos datos, útiles para la competición, plantean interrogantes en relación a la rigurosidad del registro, la fiabilidad de la herramienta utilizada o la del observador que los ha recogido (Anguera, Blanco-Villaseñor, Losada, y Hernández-Mendo, 2000), desembocando en una cuestión de validez y fiabilidad de los datos ofrecidos (Hughes y Bartlett, 2002). Por tanto, pueden ser cuestionados por la investigación, siendo necesario proceder en primera instancia a su validación para posteriormente ser utilizados por los investigadores.

Anguera et al. (2000) ponen de manifiesto la necesidad de realizar un adiestramiento de los observadores, para el cual existen diversas propuestas como la de Medina y Delgado, (1999) o la de Losada y Manolov (2014). Tras el entrenamiento, los observadores son sometidos a una evaluación utilizando coeficientes de concordancia o asociación. Se asume la fiabilidad de los datos con un grado de acuerdo (inter e intra observador) de .70 o superior (Randolph, 2005). Pero los índices de concordancia únicamente miden la consistencia en el grado de acuerdo, aspecto que sólo constituye una de las tres formas de entender la fiabilidad (Anguera et al. 2000; García, Argudo, y Alonso, 2007; Hernández-Mendo, 1996). Blanco-Villaseñor (1989, 1997) considera que existen tres formas de estimar la fiabilidad de los datos observacionales: a) mediante los coeficientes de concordancia entre dos observadores independientes que utilizan el mismo instrumento, b) a través de los coeficientes de acuerdo basados en correlaciones, c) aplicando la *Teoría de la Generalizabilidad (TG)*, en referencia a la precisión de la generalización, y por consiguiente, a la calidad del dato. El concepto de *calidad del dato*, además de integrar el concepto de fiabilidad, incluye el error del instrumento de medida, concepto vinculado al de validez, así como al de precisión de la medida, es decir, si se

representa la conducta en cuestión (Blanco-Villaseñor, 1997).

Dentro la *Teoría de la Generalizabilidad* Cronbach, Gleser, Nanda, y Rajaratnam (1972), consideran que en cualquier situación de medida existen múltiples fuentes de error. Asume que hay otras fuentes de variación además de las clásicas (inter, intra y error), e integra cada una de ellas en una estructura global. El eje central se encuentra en los componentes de la varianza, dado que su magnitud aporta información sobre las fuentes de error que están afectando una medición, mediante la estimación del grado de generalización de un diseño de medida con condiciones particulares, a un valor teórico buscado (Blanco-Villaseñor, 1989, 1997; Blanco-Villaseñor, Castellano, Hernández-Mendo, Rosa y Usabiaga, 2014).

En las últimas décadas la utilización de la metodología observacional en el estudio de la actividad física y deportiva en diversas disciplinas deportivas ha tenido un crecimiento constante y sostenido, y se detecta una solidez en el planteamiento que hace presagiar que su utilización seguirá incrementándose (Anguera y Hernández-Mendo, 2013, 2014). Esta metodología se adapta perfectamente para resolver los problemas de estudio que surgen en la educación física y el deporte.

Se han identificado diversos estudios que utilizan la observación como método para el análisis del juego. En una búsqueda en la WOK (Web Of Knowledge) y base de datos TESEO, con las palabras clave: notational analysis, reliability, validity, observational methodology, y Theory of Generalizability se hallaron diversos artículos pudiendo clasificarlos en tres grandes grupos (Tabla 1): los que cumplen todos los criterios anteriormente definidos de validez, fiabilidad y precisión de la medida (1º grupo); los estudios que incluyen la consecución de la fiabilidad a través del entrenamiento de sus observadores (2º grupo); y finalmente los que no mencionan ni definen algún tipo de control de calidad del dato (3º grupo).

**Tabla 1.** Clasificación de estudios observacionales en función del procedimiento utilizado.

Deportes	1º Grupo	2º Grupo	3º Grupo
Fútbol	Blanco-Villaseñor, Castellano, y Hernández-Mendo (2000). Blanco-Villaseñor y Oliva-Millán (2010). Castellano, Hernández-Mendo, Gómez, Fontetxa, y Bueno (2000). Castellano, Perea, y Hernández-Mendo (2008). Sarmiento, Anguera, Campanico, y Leitao (2010).		Taylor, Mellalieu, James, y Shearer (2008). Vales, Areces, Blanco, y Arce (2011)
Waterpolo	García et al. (2007)		
Tenis/ Frontenis	Gorospe, Hernández-Mendo, Anguera, y Martínez de Santos (2005). Garay, Hernández-Mendo, Morales, (2006).	Alonso y Argudo (2008)	
Baloncesto	Garzón, Lapresa, Anguera, y Arana (2011). Sautu (2010)	Ibáñez, García, Feu, Lorenzo, y Sampaio, (2009). Ortega, Fernández, Ubal, Lorenzo, y Sampaio (2010)	García et al. (2011). Martínez y Martínez, (2010)
Voleibol	Hernández-Mendo, Montoro, Reina, y Fernández (2012)	Ureña, Vavassori, León, y González (2011)	
Balonmano			Meletakos, Vagenas, y Bayios (2011). Ohnjec et al. (2008). Volossovitch et al. (2012)

Hughes, Cooper, y Nevill, (2004) ponen de manifiesto las limitaciones en el diseño de las investigaciones basadas en el análisis del rendimiento. Realizaron una revisión sistemática de trabajos presentados en conferencias internacionales, encontrando un 67% que no mencionaban la inclusión de cualquier evaluación de la fiabilidad de los datos, y un 15% que empleaba técnicas de análisis estadístico inapropiadas. Estas omisiones o incoherencias parecen ser un descuido fundamental por parte de los autores, pues son de vital importancia para la relevancia de los resultados de investigación, pudiendo argumentar así que la calidad científica de algunos trabajos de este campo no ha sido alta (Drust, 2010; Sarmiento et al., 2010).

Por lo tanto, tras la revisión realizada, el objetivo principal de este estudio es validar los indicadores de rendimiento ofrecidos por las estadísticas de juego de la liga ASOBAL. De este objetivo principal, se derivan tres objetivos secundarios: i) Evaluar la concordancia entre los observadores oficiales de la competición y los entrenados; ii) Analizar la validez de los datos registrados en las “estadísticas de juego” de la liga ASOBAL; iii) Calcular la estimación de la muestra a partir de cual se puede generalizar que los datos se pueden utilizar para la investigación.

## Método

### Taxonomía del diseño observacional

El diseño observacional de esta investigación es *nomotético, de seguimiento y multidimensional (NSM – cuadrante IV)* pues se estudia una pluralidad de unidades tratadas de forma independiente, la estabilidad del comportamiento en sucesivas sesiones, existiendo heterogeneidad en cuanto a las posibilidades de los niveles de respuesta (Anguera, Blanco-Villaseñor, y Losada, 2001; Anguera, Blanco-Villaseñor, Hernández-Mendo y Losada, 2011), la métrica del registro se basa en el parámetro *Orden*, (Anguera et al. 2001), y el tipo de dato son *Multieventos* (Bakeman y Quera, 1995).

### Participantes

Se realizó un *muestreo de eventos* de carácter focal (Anguera, 1990; Hernández-Mendo, 1996). El número de partidos analizados son el resultado del *estudio de decisión (D)*, ya que los coeficientes obtenidos indican que esa cantidad permite

generalizar los resultados. Con ello se pretende utilizar para el análisis el mínimo número de partidos que nos permita afirmar que los datos obtenidos en relación a la fiabilidad, puedan ser atribuibles a toda la temporada objeto de estudio. Esta muestra quedó definida en 3 partidos, de los que se recogieron un total de 7950 unidades de observación. Estos partidos pertenecen a las jornadas 16 y 17 de la liga ASOBAL, temporada 2011/12. La elección de los partidos se ha realizado al azar, asegurando la estabilidad de las sesiones a nivel interesional e intrasacional (Castellano et al., 2008), es decir, utilizar partidos de la misma liga y temporada, y eliminar los partidos donde se rompa la constancia a causa de un evento inesperado o circunstancia sobrevenida en su transcurso. En ningún caso la Inobservabilidad superó el 10% del total (Anguera, 1990; Anguera et al., 2000). Los indicadores de rendimientos correspondientes a la muestra (estadísticas de juego), han sido extraídos de [www.asobal.es](http://www.asobal.es).

### Variables del estudio

Las variables de estudio hacen referencia a dos aspectos. Por un lado, a las categorías del sistema de formatos de campo, provenientes de los indicadores de rendimiento que se recogen por la organización del campeonato durante todos los partidos (estadísticas de juego): localización, zona/tipo de lanzamiento, forma de finalización del lanzamiento, sanciones disciplinarias, faltas, acciones ofensivas, acciones defensivas, zonas de la portería y eficacia del portero. Las definiciones y descripciones de los indicadores fueron facilitadas por la propia organización para utilizar un sistema común. Por otro lado, a los diferentes observadores que participan en el estudio para validar el registro oficial de la competición, con sus registros: el observador oficial (OBS.1) y los entrenados (OBS.2A y OBS.2B).

Instrumento de registro: Sistema taxonómico

Para el registro de las conductas se utilizó un sistema de formatos de campo compuesto por 10 criterios y 35 unidades de registro (Tabla 2). En el sistema de formatos de campo, se denomina “criterios” a las categorías de estudio. Los criterios corresponden a los indicadores de rendimiento (estadísticas de juego) que se recogen durante la competición. En las unidades de registro se incluyen las posibilidades de registro de cada criterio.

**Tabla 2.** Sistema taxonómico: (Sistema de Formato de Campo).

Nº/Criterios	Código	Unidades de observación	Código
1-Localización	C1-LOC	Local	LC
		Visitante	VI
2-Zona/tipo de lanzamiento	C2-ZTL	Lanzamiento-contraataque	LC
		Lanzamiento-6m	L6
		Lanzamiento-9 m	L9
		Lanzamiento-7m	L7
3-Forma de finalización del lanzamiento	C3-FFL	Gol	G
		Parada	P
		Poste	PO
		Fuera	F
4-Sanciones disciplinarias	C4-SD	Amonestación	A
		Primera-Exclusión	1E
		Segunda-Exclusión	2E
		Descalificación	D
		Descalificación-directa	DD
5-Faltas	C5-F	Técnica	FT
		Reglamentaria	FR
6-Acciones ataque	C6-AA	Saque-centro inicio	SA
		Pasivo	PAS
		Tiempo muerto equipo	TME
		Pérdida-balón	PB
7-Acciones defensa	C7-AD	Lanzamiento-bloqueo	LB
		7metros-cometidos	7mC
		Recuperación-balón	REC
8-Zonas de portería	C8-ZP	Esquina-superior-izquierda	Z1
		Centro-superior	Z2
		Esquina-superior-derecha	Z3
		Centro-medio-izquierda	Z4
		Centro-medio	Z5
		Centro-medio-derecha	Z6
		Esquina-inferior-izquierda	Z7
		Centro-inferior	Z8
		Esquina-inferior-derecha	Z9
9-Eficacia portero	C9-EP	Lanzamiento-parado	PA
		Gol-encajado	GE
10-Instrumental	C10-IT	Error-omisión	EO
		Error complementario-omisión	ECO
		Error-comisión	EC

Instrumentos de observación: el observador

Se considera a los observadores entrenados como instrumentos del estudio debido a que son la herramienta que se utiliza para recoger los datos a través del instrumento de registro descrito anteriormente. Y como un instrumento más, es sometido a una serie de procedimientos. Por lo tanto, después la selección de observadores según las recomendaciones de Arias, Argudo, y Alonso (2009), se llevó a cabo el entrenamiento de observadores

(Medina y Delgado, 1999), con una duración de 12 sesiones. Tras el adiestramiento, los observadores fueron sometidos a una evaluación, donde, según la clasificación de Landis y Koch (1977), la fuerza de asociación obtenida fue *Almost Perfect* (.93 para la fiabilidad inter-observador y, .92 y .94 para la fiabilidad intra-observador) al excluir la categoría 8 *Zonas de la portería*, y una fuerza de asociación *Sustancial* (.73 para la fiabilidad inter-observador y, .80 y .72 para la fiabilidad intra-observador) al incluir la categoría en cuestión. En todos los casos la asociación fue estadísticamente significativa.

### Material

Se ha utilizado el programa SPSS 19.0 para el análisis de concordancia y fiabilidad, y el programa SAGT versión 1.0 (Hernández-Mendo, Ramos-Pérez y Pastrana, 2012) para el análisis de generalizabilidad.

### Procedimiento

Tras la construcción de la herramienta observacional, las pautas para el control de sesgos (Behar y Riba, 1993), la selección de observadores, y la planificación y desarrollo del entrenamiento de observadores, se realizó una evaluación de los mismos como paso previo al registro de los datos (primera etapa). Y una vez superados los mínimos de fiabilidad ( $K=.70$ ) establecidos por Randolph (2005), se procedió a la segunda etapa del proceso, donde los observadores entrenados visualizaron y codificaron los partidos seleccionados de forma independiente, registrando una sesión por día para evitar así algún tipo de fatiga. Los datos se registraron sobre grabaciones, de forma sistemática y continua en función del parámetro orden (O).

Una vez finalizada esta segunda etapa, se efectuó el tratamiento de los datos y los diferentes análisis en función de los objetivos propuestos.

### Análisis de datos

Estudio de concordancia inter-observadores (entrenados-oficial)

Se analizó la fiabilidad de los indicadores de rendimiento mediante el índice *Kappa de Cohen* (Cohen, 1968), y la clasificación seguida para el establecimiento de la fuerza de asociación fue la de Landis y Koch (1977), siempre y cuando fuese estadísticamente significativa ( $p<.05$ ).

Estudio de generalizabilidad (G) y estudio de decisión (D).

El control de calidad del dato se realizó en dos fases según los criterios descritos por Blanco-Villaseñor (1989, 1997): el *Estudio de generalizabilidad* (G), mediante el análisis de la varianza y los coeficientes globales de generalizabilidad, y el *estudio de*

decisión (D), para optimizar los datos obtenidos previamente, es decir, conocer si introduciendo alguna modificación en las dimensiones analizadas en el estudio, mejorarían los resultados; si esto fuese así, se debería valorar el costo-beneficio y decidir si volver a repetir el trabajo incluyendo esas modificaciones.

## Resultados

### Estudio de la concordancia inter-observador

En la tabla 3 se muestran los valores de los índices *Kappa* para las categorías valorando la fiabilidad inter-observador entre los dos sujetos entrenados así como la de cada uno de estos con el observador oficial.

Los resultados descubren un alto índice de concordancia (entre .80 y 1.00) para la mayoría de las categorías, con una fuerza de asociación *Almost Perfect* según la clasificación de Landis y Koch (1977). Se observan algunas categorías con un índice aceptable (entre .70 y .80), estableciéndose una asociación *Substantial* (Landis y Koch, 1977). Como excepción, la categoría 8 Zonas de la portería, que obtiene un valor *Kappa* claramente inferior (.58 y .57) en la concordancia inter-observador entre los observadores entrenados y el oficial, alcanzándose una fuerza de asociación *Moderate* (Landis y Koch, 1977). Asimismo, se puede advertir como, en general, los valores *Kappa* son mayores para la concordancia inter-observador entre los observadores entrenados, que para cada uno de éstos con el observador oficial.

Tabla 3. Estudio de concordancia inter-observador (por categorías).

Categorías	K: OBS.2-A/OBS.2-B	K: OBS.2-A/OBS.1	K: OBS.2-B/OBS. 1
C1-LOC	.99	.78	.78
C2-ZTL	.94	.76	.72
C3-FFL	1.00	.97	.96
C4-SD	.99	.94	.93
C5-F	1.00	.88	.88
C6-AA	.98	.83	.81
C7-AD	.98	.76	.76
C8-ZP	.90	.58	.57
C9-EP	1.00	.96	.96

Nota: OBS.1=Observador oficial; OBS.2-A=Observador entrenado A; OBS.2-B=Observador entrenado B; LOC=Localización; ZTL=Zona/tipo lanzamiento; FFL=Forma finalización del lanzamiento; SD=Sanciones disciplinarias; F=Faltas; AA=Acciones ataque; AD=Acciones defensa; ZP=Zonas de portería; EP=Eficacia del portero.

En la tabla 4 se observan los índices de concordancia de las secuencias globales, considerando la inclusión de la categoría 8, Zonas de la portería. La fiabilidad entre los observadores entrenados se mantiene por encima de .80 en ambos casos, obteniendo una fuerza de asociación *Almost Perfect* (Landis y Koch, 1977). Ocurre lo contrario cuando se trata de la con-

cordancia entre los sujetos entrenados y el observador oficial, pues, si se considera la categoría 8, el coeficiente *Kappa* obtiene valores de .48 y .46, catalogados como *Moderate*, pero si se obvia, el índice aumenta hasta .73 y .70, siendo *Substantial* (Landis y Koch, 1977).

Tabla 4. Estudio de concordancia inter-observador de la secuencia completa.

Condición	K: OBS.2A/OBS.2B	K: OBS.2A/OBS.1	K: OBS.2A/OBS.1
Incluyendo la categoría-8	.87	.48	.46
Excluyendo la categoría-8	.95	.73	.70

Debido a los resultados obtenidos para la categoría 8, Zonas de la portería, se decidió excluir para el estudio de Generalizabilidad y por extensión del estudio de Decisión.

### Estudio de Generalizabilidad (G)

Para llevar a cabo el estudio G, se estableció un *Plan de Observación* con una estructura de diseños multifaceta cruzadas,

identificando como facetas: *Observadores* (O, con 3 niveles); *Criterios* (C, 8 niveles), y *Partidos* (P, 3 niveles.). Para ello se ha diseñado tres *Planes de Medida* de dos y tres facetas. Un primer diseño de dos facetas en la que estudia el comportamiento de los dos observadores entrenados en relación a los criterios que conforman el sistema de formatos de campo, utilizando los datos de un solo partido. Un segundo diseño de dos facetas en la que se añade al tercer observador, el ofi-

cial, para ver el comportamiento de éste respecto a los datos anteriores. Con esto se pretende comprobar si la variabilidad se atribuye a los observadores, lo cual sería un indicador para pensar en un nuevo planteamiento respecto a los mismos (re-entrenamiento, número de entrenadores) o si se atribuye a los criterios, lo cual puede indicar que hay adecuada discriminación, y se podría proseguir con el estudio. Y finalmente, un tercer diseño de tres facetas planteado para estudiar la interacción conjunta de todas las facetas propuestas, y observar cómo se distribuye la variabilidad y los resultados de los coeficientes.

Por lo tanto, para la estimación de la homogeneidad y del

ajuste del sistema de observación se realizó a través de un diseño de dos facetas (Observadores/Criterios= O/C). El análisis de la varianza (Tabla 5) muestra que la totalidad de la variabilidad se encuentra asociada a la faceta *Criterios* (100%). Los coeficientes globales de generalizabilidad obtenidos (Tabla 6) fueron excelentes (.000 y .000), pues confirman tanto la adecuación del sistema como la homogeneidad de los criterios, es decir, que la bondad de los criterios es altamente significativa en el sentido de diferenciadora. Esto se traduce en que las categorías describen de forma heterogénea la realidad del juego, aspecto relacionado con la validez del sistema de observación y registro utilizado.

**Tabla 5.** Análisis de la varianza: estimación de diseños de dos y tres facetas.

FV	SSq	D.F.	M.Sq.	Random. Comp.	Mix. mod. Comp	Corrected. Comp.	Stand. error	%
<i>Diseños de dos facetas: Observadores (entrenados)*Criterios</i>								
O	0	1	0	-0.02	-0.02	-0.02	0.02	0
C	87205	7	12457.86	6228.86	6228.86	6228.86	2936.35	100
OC	1	7	0.143	0.143	0.143	0.143	0.067	0
<i>Diseño de dos facetas: Observadores , Criterios y su interacción</i>								
O	75	2	37.50	3.57	3.57	3.57	3.34	0.06
C	127664.50	7	18237.79	6076.29	6076.29	6076.29	2865.79	99.80
OC	125	14	8.93	8.93	8.93	8.93	3.16	0.15
<i>Diseños de tres facetas: Observadores*Criterios*Partido</i>								
O	160.44	2	80.22	1.73	1.73	1.73	2.45	0.03
P	709.78	2	354.89	10.55	10.55	10.55	10.55	0.16
OP	90.89	4	22.72	1.92	1.92	1.92	1.62	0.03
C	408715.06	7	58387.87	6476.16	6476.16	6476.16	3058.26	99.19
OC	328.44	14	23.46	5.36	5.36	5.36	2.84	0.08
PC	1209.11	14	86.37	26.33	26.33	26.33	10.13	0.40
OPC	206.89	28	7.39	7.39	7.39	7.39	1.91	0.11

*Nota:* O=Observadores; C=Criterios; P=Partidos; OC=Interacción Observadores-Criterios; OP=Interacción Observadores-Partidos; PC= Interacción Partidos-Criterios; OPC= Interacción Observadores-Partidos-Criterios.

Mediante un diseño de dos facetas (*Criterios/Observadores = C/O*), se calculó la fiabilidad inter-observadores. La determinación de las fuentes de varianza (Tabla. 5) reveló que la mayor parte de la variabilidad (99,80%) estaba asociada a la faceta *Criterios* mientras que fue prácticamente nula para la faceta *Observadores* (0,06%) y para la interacción *Observadores/Categorías* o *Residual* (0,15%). En el análisis global de los coeficientes de generalizabilidad (Tabla.6) se obtuvieron valores muy satisfactorios (1 y .999), demostrando la elevada fiabilidad entre los tres observadores.

**Tabla 6.** Coeficientes de generalizabilidad para diseños de dos y tres facetas.

Diseños	Modelos	Coeficientes	
		FD/FI	$\phi$ $\rho^2$
Homogeneidad/ajuste de los criterios	O/C <sub>EO.1</sub>	.000	.000
Fiabilidad inter-observador	C/O	1.000	.999
Número mínimo de partidos	OC/P	.998	.998

*Nota:*  $\phi$ =coeficiente relativo de generalizabilidad;  $\rho^2$ =coeficiente absoluto de generalizabilidad.

Finalmente, para valorar el número mínimo de partidos necesario para la consecución de la generalización con precisión, se utilizó un diseño de tres facetas (*Observadores-Criterios/Partidos=OC/P*). La determinación de las fuentes de varian-

za reveló que la mayor parte de la variabilidad se asociaba a la faceta *Criterios* (99,19%), siendo mínima o prácticamente inexistente para el resto de facetas e interacciones (Tabla 5). En el análisis global de los coeficientes de generalizabilidad (Tabla. 6) se alcanzaron valores muy satisfactorios (.998).

### Estudio de Decisión

Para optimizar los *Planes de Medida*, se efectuó en terminología de Cronbach et al. (1972) un estudio de decisión. En este caso, fueron utilizadas de forma conjunta las tres facetas propuestas, de manera que cada una de ellas actuó como faceta instrumental. Cada vez que una de estas facetas obtuvo tal rol, se sometió a modificaciones en cuanto al número de niveles, obteniendo unos coeficientes de generalización estimados (Blanco-Villaseñor, 1989, 1997). (pe. para el diseño OC/P, donde P=Partido actúa como faceta instrumental, se estimó los coeficientes con un valor P diferentes al estudio real).

Para el diseño (PC/O) se propusieron tres modelos variando el número de niveles de la faceta *Observadores* (Tabla 7). Los resultados muestran que al aumentar el número de niveles de esta faceta los valores de los coeficientes (.999) se mantienen similares al obtenido en el estudio G (.999), únicamente si se utilizaran seis o más observadores, como se observa en la opción 2(OP-2), el modelo sería perfecto (1.000).

En el diseño (OP/C) se modificó el número de niveles de la faceta *Criterios*. Los resultados mostraron que al aumentar el número de criterios no mejoraban los coeficientes obtenidos (OP-1=.845 y .032), al contrario de lo que ocurrió al disminuir esta cantidad (OP-2= .592 y .009).

Finalmente, en el diseño (OC/P), los modelos planteados ( $P=6$  y  $P=15$ ) obtuvieron unos coeficientes muy próximos a los conseguidos en el estudio G realizado previamente. En la opción 2 (Op-2) se lograron unos coeficientes perfectos (1.000).

Tabla 7. Estudio de decisión: optimización de los diseños de tres facetas y coeficientes de generalizabilidad.

Facetas	Fiabilidad Inter-observ.			Homogeneidad y ajuste			Nº mínimo de partidos		
	PC/O			OP/C			OC/P		
	E-G	OP-1	OP-2	E-G	OP-1	OP-2	E-G	OP-1	OP-2
<i>Observadores</i>	3; INF	4; INF	6; INF	3; INF	3; INF	3; INF	3; INF	3; INF	3; INF
<i>Partidos</i>	3; INF	3; INF	3; INF	3; INF	3; INF	3; INF	3; INF	6; INF	15; INF
<i>Criterios</i>	8; INF	8; INF	8; INF	8; INF	15; INF	4; INF	8; INF	8; INF	8; INF
Coefficientes G	E-G	OP-1	OP-2	E-G	OP-1	OP-2	E-G	OP-1	OP-2
<i>Relativo (<math>\phi</math>)</i>	.999	.999	1.000	.744	.845	.592	.998	.999	1.000
<i>Absoluto (<math>\rho_2</math>)</i>	.999	.999	1.000	.017	.032	.009	.998	.999	1.000

Nota: OP=Opción; E-G=Estudio de generalizabilidad;

Los resultados ponen de manifiesto que para la muestra analizada, el proceso de validación de los indicadores de rendimiento puede realizarse con 3 partidos, pudiendo generalizar la validación a toda la muestra de datos de una temporada.

Aunque en este estudio no sería necesario llevar a cabo el estudio de decisión con el objetivo de optimizar los coeficientes de generalizabilidad (pues se puede observar que ya fueron adecuados), si se puede realizar con el objetivo de comprobar y corroborar que una mayor inversión temporal, económica o de recursos, no mejoraría sustancialmente los resultados. Desde otro punto de vista, con este estudio de decisión se demuestra la necesidad de estimar, mediante el mismo, el número de niveles de cada una de las facetas que cualquier investigador se proponga (número de observadores, de partidos o de criterios del sistema de formato de campo, en este caso), porque de este modo puede predecir los recursos necesarios y la gestión de los mismos, para obtener los mejores resultados antes de llevar a cabo el estudio.

### Discusión

La importancia de este trabajo radica en la presentación de un procedimiento para estipular, como investigadores, la calidad de los datos observacionales obtenidos en los medios de difusión deportivos. Estos datos son usados diariamente por equipos profesionales y empleados por los investigadores para realizar su investigación. Si los datos que provienen de las competiciones oficiales son adecuados a la realidad de juego, éstos podrán ser utilizados para la investigación con total garantía.

Para responder al objetivo principal del estudio, se han seguido los pasos propuestos en la literatura, tales como Anguera et al. (2000), Anguera et al. (2011), para el procedimiento metodológico y el diseño de investigación utilizado, Blanco-Villaseñor y Oliva-Millán (2010), para el estudio de fiabilidad y generalización en el estudio de conductas a través de la observación. Y finalmente Castellano et al. (2008), Hernández-Mendo et al. (2012), Hernández-Mendo, Díaz-Martínez, y Morales-Sánchez (2010) y Reina-Gómez, Her-

nández-Mendo y Fernández-García (2009) donde se hallan diversas aplicaciones de una o varias fases del proceso de validación de datos en diferentes modalidades deportivas. El procedimiento llevado a cabo se ajusta a los parámetros establecidos por los autores de referencia en este ámbito.

En cuanto a los objetivos específicos, los resultados del primer estudio, análisis de la concordancia entre observadores en el registro de los indicadores de rendimiento, han sido concluyentes. Se observa un grupo de categorías con resultados óptimos, pues el índice de fiabilidad ha sido casi perfecto (Landis y Koch, 1977). En cuanto al resto de categorías la fiabilidad entre los observadores se mantiene elevada y estadísticamente significativa aunque su fuerza de asociación es menor. Las *Zonas de la portería*, (categoría 8) donde se consigue el gol, es la variable más difícil de registrar.

Estos resultados pueden ser consecuencia de diversos aspectos no considerados en la planificación, dando lugar, tal y como exponen Behar y Riba (1993) a sesgos que afectan al observador y a la observación. Por un lado pueden deberse a los diferentes contextos de observación del observador oficial respecto a los observadores entrenados que analizan las acciones en condiciones óptimas. El primero, desarrolla su labor en el campo de juego (directo o in vivo), mientras que los segundos lo hacen en laboratorio (indirecto) y a través de grabaciones, donde, además de la diferencia en el tipo de registro y en el entorno en el que se realiza, probablemente no coincida el lugar de grabación con la situación del observador oficial. En este sentido, Losada y Arnau (2000) afirman que en el uso de medios audiovisuales para la observación, la zona de grabación puede llegar a influir para el registro de determinadas conductas o acciones ofreciendo ciertos problemas de inobservabilidad. En la misma línea, Anguera et al. (2000) establecen un tipo de sesgo que amenaza negativamente a la observación, denominándolos *fallos de procedimiento*, entre los que incluye los ángulos de visión incorrectos o el uso de diferentes ángulos de observación por parte de observadores estudiados. Behar y Riba (1993) se refieren a ellos como errores del observador que no producen reactividad, y los definen como sesgos de tipo perceptivo. Entre ellos se destacan: la *localización espacial*, la *duración del estímulo*, o la *selección de la atención*. Asimismo, Behar y Riba (1993) definen los *efectos situacionales* como sesgos que producen reactividad, es decir, que las condiciones bajo las que se lleva a cabo la observación (alto grado de actividad, elevados niveles de ruido o alta frecuencia de movimiento) pueden llegar a influir en la tarea desarrollada por el observador. Por otro lado, estos resultados también pueden venir motivados por el sistema de observación elaborado para el estudio, pues no es utilizado por el observador evaluado. Por esta causa, se pueden encontrar discrepancias a la hora de categorizar determinadas conductas, influyendo en lo que Anguera et al. (2000) denominan como errores de omisión o comisión.

En cuanto al estudio de concordancia global entre los dos observadores entrenados y el oficial, los índices de fiabilidad obtenidos al excluir la variable 8 del estudio han sido aceptables, estableciendo así que los indicadores de rendimiento son fiables (Landis y Koch, 1977). De esta manera, la variable 8 *Zonas de la portería*, quedó eliminada del estudio, pues se vio notablemente afectada por ciertos sesgos que imposibilitaron el estudio de su fiabilidad.

El segundo objetivo específico del estudio, análisis de la validez de los indicadores de rendimiento, y en referencia al estudio G, los resultados del análisis de la varianza mostraron que, el mayor porcentaje de variabilidad se asoció a la faceta *Criterio* (indicadores de rendimiento). Resultados similares han sido obtenidos en los trabajos de Blanco-Villaseñor et al. (2000), Castellano et al. (2000), Gorospe et al. (2005), Hernández-Mendo et al. (2010), y Sautu (2010), de los que se deduce la necesidad de que la principal fuente de variación sea la faceta *Criterios* (para ellos *Categorías*), pues se confirma que los criterios propuestos discretizan las acciones o elementos a observar. Los resultados ponen de manifiesto la importancia de una correcta definición y delimitación de los indicadores de rendimiento, acciones de juego, categorías o estadísticas de juego, para así evitar situaciones de incertidumbre o confusión a la hora de incluir una acción en una u otra categoría.

Por el contrario, es un buen indicador que no se asocie excesiva variabilidad a las facetas *Observadores* y *Partidos*, pues hacen referencia a la fiabilidad de los observadores y a la constancia inter-sesional respectivamente.

En cuanto a los coeficientes globales de generalizabilidad obtenidos y según las bases de esta teoría (Cronbach et al., 1972), los resultados han sido adecuados, ofreciéndonos la siguiente información:

Mediante el diseño *Criterios/Observadores* determina una excelente fiabilidad de precisión de los resultados en relación a la fiabilidad inter-observador, al igual que reflejan en sus estudios Blanco-Villaseñor et al. (2000), y Castellano et al. (2000) en fútbol; Gorospe et al. (2005) en tenis; Hernández-Mendo et al. (2012) en voleibol; y Sautu (2010) en baloncesto.

El estudio del diseño *Observadores/Criterios* pone de manifiesto la validez del sistema de formatos de campo utilizado para la observación y registro de los datos. Resultados análogos se muestran en los trabajos de Gorospe et al. (2005), Hernández-Mendo et al. (2012), y Sautu (2010), en los que afirman que estos valores permiten certificar que el sistema de observación propuesto (para ellos sistema de categorías) diferencia adecuadamente las acciones que se registran en cada una de las categorías, apuntando que la bondad de los criterios establecidos resulta altamente significativa, en el sentido de diferenciadores.

Por último, a través del diseño *Observadores-Criterios/Partidos*, se establece una alta precisión en la generalización de los resultados con el número de partidos analizados. Como

indican Gorospe et al. (2005), Hernández-Mendo et al. (2010), Hernández-Mendo et al. (2012), y Sautu (2010), el número de partidos analizados han sido suficientes para generalizar los resultados de la muestra estudiada a la población universo.

Finalmente, en cuanto al tercer objetivo específico de la investigación, los resultados del estudio de *Decisión (D)* para el diseño OC/P, confirman, según las bases de la *TG* (Cronbach et al., 1972), la adecuación del modelo establecido, pues si se aumentara el número de partidos a analizar, los coeficientes se mantendría similares. Estos resultados van en consonancia con los obtenidos por Gorospe et al. (2005), al obtener unos resultados adecuados en el estudio *G*, las propuestas para optimizar los *Planes de Medida* no serían eficaces en cuanto a costo/beneficio. Por el contrario, en el trabajo de Blanco-Villaseñor et al. (2000) les sirvió para mejorar considerablemente sus aportaciones.

Como se ha mostrado, diversos autores han utilizado esta teoría para sus trabajos de investigación, pues ofrece un marco satisfactorio para buscar las estimaciones de precisión y generalización de los márgenes de error, debido a que es suficientemente globalizada como para adaptarse a las condiciones particulares de cada objeto de medida (Blanco-Villaseñor, 1997), aspecto por el que también se impide la comparación de estudios realizados en diferentes condiciones de observación, aún definiendo similares fuentes de variación.

En definitiva, este tipo de estudios sitúa al investigador en una posición inmejorable para tomar decisiones, permitiéndole desarrollar la investigación en las mejores condiciones (Castellano et al., 2008).

A pesar de que el balonmano es un deporte de alta complejidad a la hora de recoger información de forma directa, los indicadores de rendimiento registrados por los observadores oficiales de ASOBAL que han sido estudiados, han cumplido los requisitos de fiabilidad exigidos por el riguroso proceso de validación propuesto. Además, estos resultados pueden generalizarse de forma precisa a la población que conforma este estudio. Por lo tanto, los datos validados que provienen de los registros oficiales de la competición ASOBAL pueden ser utilizados para desarrollar una investigación a mayor escala.

### Aplicaciones prácticas

La validación de los indicadores de rendimiento, acciones de juego que provienen de las estadísticas de juego, es fundamental para el desarrollo de investigaciones sobre el análisis de la competición. El empleo de la *TG* para este procedimiento permite reducir la muestra sobre la que trabajar durante el proceso de validación, reduciendo los tiempos empleados en esta fase de la investigación. Los investigadores deben familiarizarse con esta técnica e incluirla dentro de sus estudios.

Las investigaciones que se realizan sobre el análisis de la competición, deben incluir dentro de su proceso metodológico la validación de los datos, así como el entrenamiento de los codificadores encargados de la recopilación de los datos. A partir de estos datos, los investigadores podrán establecer las hipótesis causales (diferencias), de covariación (relaciones), o predicciones en los indicadores de rendimiento en cada uno de los deportes analizados.

### Referencias

- Alonso, J. I. y Argudo, F. (2008). Performance indicators for olympic frontenis female serve. *International Journal of Sport Science*, 4(10), 59-76.
- Anguera, M. T. (1990). Metodología observacional. En J. Arnau, M. T. Anguera y J. Gómez (Eds.), *Metodología de la investigación en Ciencias del Comportamiento* (pp.125-236). Murcia: Secretariado de Publicaciones Universidad de Murcia.
- Anguera, M. T. y Hernández-Mendo, A. (2013). Metodología observacional en el ámbito del deporte. *e-balonmano.com: Revista de Ciencias del Deporte*, 9 (3), 135-160.
- Anguera, M. T. y Hernández-Mendo, A. (2014). Metodología observacional y psicología del deporte: Estado de la cuestión. *Revista de Psicología del Deporte*, 23, (1), 103-109.
- 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., 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., Losada, J. L. y Hernández-Mendo, A. (2000). La metodología observacional en el deporte: Conceptos básicos. *Lecturas: E.F. y Deportes, Revista Digital*, 24. Recuperado de <http://www.efdeportes.com/efd24b/obs.htm>
- Arias, J. L., Argudo, F. M. y Alonso, J. I. (2009). El proceso de formación de observadores y la obtención de fiabilidad en metodología observacional para analizar la dinámica de juego en minibasket. *Apunts*, 98, 40-45.
- Bakeman, R. y Quera, V. (1995). *Analyzing interaction: Sequential analysis with SDIS and GSEQ*. New York: Cambridge University Press.
- Behar, J. y Riba, C. (1993). Sesgos del observador y de la observación. En M.T. Anguera (Ed.), *Metodología observacional en la investigación psicológica: Fundamentación (Vol. 2)* (pp. 15-128). Barcelona: PPU.
- Blanco-Villaseñor, A. (1989). Fiabilidad y generalización de la observación conductual. *Anuario de Psicología*, 43(4), 5-32.
- Blanco-Villaseñor, A. (1997). *Precisión en la evaluación de la Investigación Observacional*. En V Congreso de Metodología de las Ciencias Humanas y Sociales. AEMCCO, Sevilla: Septiembre.
- Blanco-Villaseñor, A. y Oliva-Millán, C. (2010). Generalización de la acción de éxito en fútbol. *Revista de Iberoamericana de Psicología del Ejercicio y el Deporte*, 5(2), 283-296.
- Blanco-Villaseñor, A., Castellano, J. y Hernández-Mendo, A. (2000). Generalizabilidad de las observaciones de la acción del juego en el fútbol. *Psicothema*, 12(2), 81-86.
- Blanco-Villaseñor, A., Castellano, J., Hernández-Mendo, A., Rosa, C. y Usabiaga, O. (2014). Aplicación de la TG en el deporte para el estudio de la fiabilidad, validez y estimación de la muestra. *Revista de Psicología del Deporte*. 23 (1), 131-137.

16. Castellano, J., Hernández-Mendo, A., Gómez de Segura, P., Fontetxa, E. y Bueno, I. (2000). Sistema de codificación y análisis de la calidad del dato en el fútbol de rendimiento. *Psicothema*, 12(4), 635-641.
17. Castellano, J., Perea, A. y Hernández-Mendo, A. (2008). Análisis de la evolución del fútbol a lo largo de los mundiales. *Psicothema*, 20(4), 928-932.
18. Cohen, J. A. (1960). Coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 20, 37-46.
19. Cronbach, L. J., Gleser, G. C., Nanda, H. y Rajaratnam, N. (1972). *The dependability of behavioral measurements: Theory of Generalizability for scores and profiles*. New York: John Wiley and Sons.
20. Drust, B. (2010). Performance analysis research: Meeting the challenge. *Journal of Sports Sciences*, 28(9), 921-922.
21. Garay, J. O., Hernández-Mendo, A., Morales, V. (2006). Sistema de codificación y análisis de la calidad del dato en el tenis de dobles. *Revista de Psicología del Deporte*, 15(2), 279-294.
22. García, J., Ibáñez, S., Parejo, I., Feu, S. y Cañadas, M. (2011). Diferencias entre nivel de juego y categoría de los jugadores en etapas de formación. *Revista Española de Educación Física y Deportes*, 395 (21), 13-28.
23. García, P., Argudo, F. y Alonso, J. I. (2007). Validación de un entrenamiento de observadores para el análisis de una microsituación de juego en waterpolo. *Lecturas: E. F. y Deportes, Revista Digital*, 109. Recuperado de <http://www.efdeportes.com/efd109/validacion-de-un-entrenamiento-de-observadores-en-waterpolo.htm>
24. Garzón, B., Lapresa, D., Anguera, M. T y Arana, J. (2011). Análisis observacional del lanzamiento de tiro libre en jugadores de baloncesto base. *Psicothema*, 23(4), 851-857.
25. Gorospe, G., Hernández-Mendo, A., Anguera, M. T. y Martínez de Santos, R. (2005). Desarrollo y optimización de una herramienta observacional en el tenis de individuales. *Psicothema*, 17(1), 123-127.
26. Hernández-Mendo, A. (1996). *Observación y análisis de patrones de juego en deportes sociomotores*. Universidad de Santiago de Compostela: Servicio de Publicaciones e Intercambio Científico.
27. Hernández-Mendo, A., Díaz-Martínez, F. y Morales-Sánchez, 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.
28. Hernández-Mendo, A., Montoro, J., Reina, A. y Fernández, 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.
29. Hernández-Mendo, A., Ramos-Pérez, F. y Pastrana, J. L. (2012). SAGT: Programa informático para análisis de Teoría de la Generalizabilidad. SAFE CREATIVE Código: 1204191501059.
30. Hughes, M. y Bartlett, R. (2002). The use of performance indicators in performance analysis. *Journal of Sports Sciences*, 20(10), 739-754.
31. Hughes, M., Cooper, S. M. y Nevill, A. (2004). Analysis of notation data: Reliability. En M. Hughes y I. M. Franks (Eds.), *Notational analysis of sport* (189-204). London: Routledge.
32. Ibáñez, S.J., García, J., Feu, S., Lorenzo, A. y Sampaio, J. (2009). Effects of consecutive basketball games on the game-related statistics that discriminate winner and losing teams. *Journal, of Sports Science and Medicine*, 8, 458-462.
33. Landis, J. R. y Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33, 159-174.
34. Losada, J. L. y Arnau, J. (2000). Fiabilidad entre observadores con datos categóricos mediante el Anova. *Psicothema*, 12(2), 335-339.
35. Losada, J. L., y Manolov, R. (2014). The process of basic training, applied training, maintaining the performance of an observer. *Quality & Quantity*, 48(2). DOI 10.1007/s11135-014-9989-7
36. Martínez, J. A. y Martínez, L. (2010). Un método probabilístico para las clasificaciones estadísticas de jugadores en baloncesto. *Revista Internacional de Ciencias del Deporte*, 18(6), 13-36.
37. Medina, J. y Delgado, M. A. (1999). Metodología de entrenamiento de observadores para investigaciones sobre E. F. y deporte en las que se utilice como método la observación. *Revista Motricidad*, 5, 69-86.
38. Meletakos, P., Vagenas, G. y Bayios, I. (2011). A multivariate assessment of offensive performance indicators in men's handball: Trends and differences in the World Championships. *International Journal of Performance Analysis in Sport*, 11(2), 284-294.
39. Ohnjec, K., Vuleta, D., Milanović, D. y Gručić, I. (2008). Performance indicators of teams at the 2003 World Handball Championship for women in Croatia. *Kinesiology* 40(1), 69-79.
40. Ortega, E., Fernández, R., Ubal, M., Lorenzo, A. y Sampaio, J. (2010). Indicadores de rendimiento defensivo en baloncesto en los ganadores y perdedores. *International Journal of Sport Sciences*, 6(19), 1-12.
41. Prieto, M. (2011). *Análisis de las acciones ofensivas que acaban en lanzamiento en el floorball de alto nivel: Comparación entre los equipos nacionales de Suecia y España*. (Tesis doctoral inédita). Universidad Politécnica de Madrid, Madrid.
42. Randolph, J. J. (October, 2005). *Free-marginal multirater kappa: An alternative to Fleiss' fixed-marginal multirater kappa*. Trabajo presentado en el Joensuu University Learning and Instruction Symposium, Finland.
43. Reina-Gómez, A., Hernández-Mendo, A. y Fernández-García, J. C. (2009). Multi-facet design for goal scoring in SOCCER-7. *Quality & Quantity*, 44(5), 1025-1035.
44. Sampaio, J., Ibáñez, S. y Lorenzo, A. (2013). Applied sports performance analysis. Basketball. En T. McGarry, P. O'Donoghue J. y Sampaio (Eds.), *Routledge Handbook of Sports Performance Analysis* (pp. 357-358). New York: Routledge.
45. Sarmento, H., Anguera, M. T., Campanico, J. y Leitao, J. (2010). Development and validation of a notational system to study the offensive process in football. *Medicina (Kaunas)*, 46(6), 401-407.
46. Sautu, L. M. (2010). *Observación y análisis de la acción del juego en baloncesto ACB*. (Tesis doctoral inédita). Universidad del País Vasco, Vitoria.
47. Taylor, J. B., Mellalieu, S. D., James, N. y Shearer, D. A. (2008). The influence of match location, quality of opposition, and match status on technical performance in professional association football. *Journal of Sports Sciences*, 26(9), 885-895.
48. Ureña, A., Vavassori, R., León, J. y González, M. (2011). Jump serve incidence on the attack phase in the Spanish under-14 volleyball. *International Journal of Sport Science*, 7, 384-392.
49. Vales, A., Areces, A., Blanco, H. y Arce, C. (2011). Diseño y aplicación de una batería multidimensional de indicadores de rendimiento para evaluar la prestación competitiva en el fútbol de alto nivel. *Journal of Sport Science*, 7(23), 103-112.
50. Volossovitch, A., Dumangane, M. y Rosati, N. (2012). Factores que influenciam a dinâmica do rendimento colectivo no andebol. *Revista Española de Educación Física y Deportes*, 396(22), 13-33.