

Revista de Psicología del Deporte  
2010. Vol. 19, núm. 2, pp. 305-318  
ISSN: 1132-239X

Universitat de les Illes Balears  
Universitat Autònoma de Barcelona

# CONSTRUCCIÓN DE UNA HERRAMIENTA OBSERVACIONAL PARA EVALUAR LAS CONDUCTAS PROSOCIALES EN LAS CLASES DE EDUCACIÓN FÍSICA<sup>1</sup>

Antonio Hernández Mendo, Francisco Díaz Martínez\*  
y Verónica Morales Sánchez\*\*

*THE CONSTRUCTION OF AN OBSERVATIONAL TOOL TO EVALUATE PRO-SOCIAL BEHAVIOURS IN PHYSICAL EDUCATION CLASSES*

KEYWORDS: Pro-social behaviour, Observation, Physical education.

ABSTRACT: This study describes the construction and validation of an original, “*ad hoc*” taxonomic system that allows the pro-social behaviours produced during compulsory secondary school physical education classes to be observed. Such behaviours are related to the development of positive values. Technically, the tool combines field format systems and E/ME category systems, which have been developed from a deductive perspective under the theoretical framework of Kohlberg’s development of moral reasoning. The tool’s optimisation process includes monitoring and analysing the quality of the data and focuses on consensus-based concordance and the analysis of the generalisability theory, which yielded good results.

Correspondencia: Antonio Hernández Mendo. Departamento de Psicología Social, Antropología Social, Trabajo Social y Servicios Sociales. Facultad de Psicología. Campus de Teatinos, s/n. Universidad de Málaga. 29071 Málaga.

E-mail: mendo@uma.es

<sup>1</sup> Este trabajo ha sido financiado por el proyecto: Avances tecnológicos y metodológicos en la automatización de estudios observacionales en deporte. (PSI2008-01179)

\* Universidad de Málaga

\*\* vomorales@uma.es

— Fecha de recepción: 19 de Junio de 2009. Fecha de aceptación: 2 de Mayo de 2010.

El desarrollo moral de los adolescentes, es una cuestión relevante para la sociedad actual y para el sistema educativo, el cual tiene entre sus funciones, la promoción de la educación integral de las personas, donde queda incluida la formación en valores y el desarrollo moral. No obstante, la crisis de valores de las sociedades occidentales, ha puesto de relieve la necesidad de buscar cambios efectivos en los programas curriculares educativos. Dichos cambios, deben de ir encaminados a otorgar a la educación moral, mayor importancia de la concedida hasta el momento. A partir de la década de los 80 y sobre todo de los 90, han venido aumentando el número de estudios científicos en éste ámbito, y más concretamente dentro del área de educación física (EF) Estos estudios, han demostrado la importancia de la EF, y el deporte, como un medio adecuado para promover el desarrollo moral en los adolescentes (Romance, 1984; Lee, 1988; Bredemeier, 1991; Escartí y García, 1994; Gibbons, Ebbeck, y Maureen, 1995; Hellinson, 1995; Weiss y Smith, 1999; Gutiérrez y Escartí, 2006).

Este trabajo se caracteriza por el abordaje de tres cuestiones metodológicas importantes. La primera está referida al uso de la Metodología Observacional para la evaluación del desarrollo positivo en las clases de educación física. Esta problemática tiene unas implicaciones epistemológicas importantes como es la diferenciación entre observación como técnica y la observación como método. La observación usada como método tiene entidad suficiente para la obtención de conocimiento científico y como técnica no supone más que un tipo de estrategia subordinado a las directrices de otra metodología (selectiva o experimental) (Anguera, 1990; Garay Plaza, Hernández Mendo y Morales Sánchez, 2006, 2007). La segunda cuestión

afronta la problemática del uso de la *Teoría de la Generalizabilidad* en los procesos de evaluación del desarrollo positivo (Blanco Villaseñor y Hernández Mendo, 1998; Blanco Villaseñor, Castellano Paulis y Hernandez Mendo, 2000).

En relación al primer aspecto de los señalados, en distintos estudios anteriores, (Boixadós, 1998, Cruz, Boixadós, Valiente y Torregrosa, 1999; Escartí, Pascual y Gutiérrez; 2005) se ha usado la observación como técnica y no como método. Esta matización es primordial ya que permite tomar decisiones metodológicas de trascendencia como es el tipo de diseño observacional (Anguera, Blanco Villaseñor, Losada, y Hernández Mendo, 2000), el tipo de datos (Bakeman y Quera, 1995), etc. La elección del tipo de diseño es de suma trascendencia ya que condicionará no solo el tipo de dato y el muestreo sino también el tipo de análisis que se realizará (Anguera, Blanco Villaseñor, Losada y Hernández Mendo, 2000). La metodología observacional ha adolecido tradicionalmente de líneas de investigación en las cuales se pusieran a prueba las múltiples posibilidades de análisis de sus datos. Probablemente, el principal motivo se halla en la superficialidad con que se obtenían tales datos, y, por consiguiente, en su carácter inconsistente. En la actualidad se ha avanzado en la configuración de los principales diseños observacionales, los cuales son de disposición no estándar en coherencia con el carácter sumamente flexible de la metodología observacional, y si además tenemos en cuenta una sustancial mejora que en la mayoría de los casos se pone en práctica en el proceso de sistematización y optimización de los datos, es lógico que hayan surgido nuevas propuestas en lo que se refiere al análisis de los mismo. Se pueden establecer criterios muy diversos en este punto de encrucijada. Desde

hace más de una década estamos trabajando con un planteamiento en donde se cruzan las *unidades de estudio*, idiográfico (unidad)/-nomotético(pluralidad), con la relativa a la *temporalidad*, una sesión vs. varias sesiones; con un tercer criterio *la dimensionalidad* (unidimensional vs. multidimensional). Esta clasificación facilita el deslinde de las direcciones básicas de análisis de datos observacionales. Con esta propuesta se dispone de cuatro cuadrantes y ocho tipos de diseño (dos por cada cuadrante), que corresponderían, a los diseños diacrónicos, sincrónicos y mixtos o *lag-log*, que absorben la totalidad de las contingencias que puedan presentarse (Anguera, 1994; Anguera, Blanco Villaseñor y Losada, 1995).

En relación a la segunda cuestión planteada, la *Teoría de la Generalizabilidad* (TG) ofrece un marco apropiado en la búsqueda de las estimaciones de la fiabilidad y de los márgenes del error. La base de la TG (Cronbach, Gleser, Nanda y Rajaratnam, 1972) se asienta en el postulado según el cual, en cualquier situación de medida existen múltiples fuentes (*facetas*) de variación. Podemos considerarla como una teoría de los errores multifaceta de una medición conductual y el objetivo es desglosar, en cualquier tipo de medición, la variabilidad real de la variabilidad del error (Blanco Villaseñor, 2001). El objetivo es analizar las diferentes fuentes de variación que pueden estar afectando a los datos. La aplicación de esta teoría permite estimar el grado de generalización de un diseño de medida con condiciones particulares a un valor teórico buscado. El coeficiente de generalizabilidad permite estimar el ajuste de la media observada a la media de todas las observaciones posibles. En este caso el problema de la *fiabilidad* se refiere a la precisión de esta generalización. (Blanco Villaseñor, 1993). El eje central de la TG se

encuentra en los componentes de varianza, dado que su magnitud nos aporta información sobre las fuentes de error que están afectando una medición conductual (Blanco Villaseñor, Castellano y Hernández Mendo, 2000; Blanco Villaseñor, 2001; Castellano y Blanco Villaseñor, 2005).

La última cuestión es relativa a la evaluación del desarrollo moral en las clases de EF. En el trabajo de Planchuelo, Fernández García y Hernández Mendo (2007) se construye una herramienta observacional para evaluar este desarrollo en los niños de primaria (Planchuelo, Hernández Mendo y Fernández García, 2009). La herramienta está formada por un sistema de formatos de campo de diez criterios y treinta y nueve categorías. Cada criterio incluye un sistema de categorías exhaustivo y mutuamente excluyente (E/ME). Los criterios se presenta de forma polar, es decir, *son necesariamente positivos o negativos* (Marías, 1941). Cada uno de los criterios está formado por un número variable de categorías. La estructura categorial se compone de un núcleo, el grado de apertura, la relación deductiva con los estadios psicológicos de Piaget y Kohlberg, el tipo de actividad motriz preferente (psicomotriz, cooperación o cooperación-oposición), y el tipo de datos predominante (eventos multimodales).

En el presente trabajo, se expone el desarrollo y optimización de un instrumento, que pretende recoger información objetiva acerca de las conductas prosociales que aparecen durante el transcurso de las sesiones de EF en secundaria. La diferencia con el trabajo de Planchuelo, Hernández Mendo y Fernández García (2009), radica tanto en el objetivo al que está orientado como las personas objeto de muestreo. Técnicamente, se trata de un sistema mixto, que combina los sistemas de formatos de campo con los

sistemas de categorías, los cuales se han desarrollado bajo el marco teórico de los postulados del desarrollo moral de Kohlberg (1981). El desarrollo y optimización de la herramienta, incluye el control y análisis de calidad del dato, requisito necesario desde la lógica científica, el cual nos permite considerar la herramienta, como fiable y precisa, siguiendo la perspectiva de diferentes trabajos (Anguera, 1979; Hernández Mendo, 1996; Ardá, 1998; Gorospe, 1999 y Castellano, 2000; Castellano, Hernández Mendo y Haro Romero, 2002; Anguera, Hernández Mendo, Gorospe y Martínez de Santos, 2005).

## Método

### Participantes

El proceso de construcción de la herramienta observacional se inició con observaciones no sistemáticas, con el objetivo de recoger información para elaborar la herramienta "ad hoc". Se estableció así un sistema provisional de formatos de campo y de sistema de categorías E/ME que se va someter a análisis. Se han observado para ello tres sesiones de EF, grabadas en DVD, correspondiente a las sesiones 4, 5 y 6 de la unidad didáctica de deportes colectivos de Voleibol, correspondientes al curso de 2º de ESO del IES "Huerta Alta", de la localidad de Alhaurín de la Torre, Málaga.

### Material

Las grabaciones se realizaron utilizando dos cámaras digitales, modelo Sony Handycam, DCR-DVD 306 E, incluyendo dos soportes situados a 1,75 cm del suelo. En el análisis de los datos se han utilizado el paquete estadístico SPSS para Windows versión 13.0, la hoja de cálculo Excel XP y el programa informático *Generalizability Study* versión 2.0.E. (Ysewijn, 1996), así como el programa

informático *SDIS-GSEQ* (Bakeman y Quera, 1995), para Windows.

### Análisis estadístico

El análisis de calidad del dato se realizó calculando la fiabilidad inter e intra observadores, para lo cual se utilizó el índice *Tau* de Kendall y la *Kappa* de Cohen. Posteriormente se complementó con un análisis de componentes de varianza a partir del cual se realizó un análisis de generalizabilidad.

### Procedimiento

La taxonomía conductual, como ya se indicó, es mixta, combinando dos estrategias diferentes en su construcción: los sistemas de categorías y los formatos de campo (Anguera, 1979; Hernández Mendo, 1996; Ardá, 1998; Gorospe, 1999 y Castellano, 2000). Los formatos de campo garantizan el registro sistemático de varios aspectos de un evento natural, para lo cual se proponen criterios relevantes y para cada uno de ellos se desarrolla un sistema de categorías situados bajo su cobertura (Anguera 1979), que hemos optado porque sean cerrados, y exhaustivos y mutuamente excluyentes. El sistema está constituido por 5 criterios y 37 categorías (ver Tabla 1).

Complementariamente, los criterios, abarcan sus correspondientes Macro categorías, las cuáles son dicotómicas en todos los criterios, exceptuando el último. En cuanto a la interacción de las conductas con las tareas, éstas pueden ser de tres tipos, individuales, de cooperación, y de cooperación-oposición. Así mismo los diferentes niveles de los estadios de desarrollo moral de Kohlberg, quedan reflejados en función de la categoría y su expansión. En la Tabla 2, presentamos la categoría *I*, a modo de ejemplo.

Las unidades de observación, corresponden a multieventos, que describen las múltiples conductas que pueden sucederse en

<b>Criterio 1. Conductas hacia sus Iguales</b>	
<i>1.1. Macro categoría de conductas positivas hacia sus iguales</i>	
A	Asiste, auxilia a los demás participantes.
I	Introduce a los participantes en la tarea.
CO	Colabora en el desarrollo de la tarea con los demás participantes.
S	Busca una mejora de nivel o perfeccionamiento con la práctica repetida.
AP	Anima a los demás participantes durante el desarrollo de la tarea.
DM	Participa con claro autocontrol en la tarea.
R	Desarrolla de forma adecuada su implicación en el aprendizaje.
<i>1.2. Macro categoría de conductas negativas hacia sus iguales</i>	
NDM	Falta de autocontrol en la tarea.
AV	Ofende con gestos o acciones a los demás participantes.
NI	No introduce a los participantes en la tarea.
NCO	No colabora en el desarrollo de la tarea, con los demás participantes.
IN	Molesta a los demás participantes durante el desarrollo de las tareas.
NR	No desarrolla de forma adecuada su implicación en el aprendizaje.
<b>Criterio 2. Conductas hacia el docente</b>	
<i>2.1. Macro categoría de conductas positivas hacia el docente</i>	
PA	Presta atención en las explicaciones del docente.
AS	Colabora con el docente en la organización de la tarea.
PAR	Participa de forma activa en la tarea.
<i>2.2. Macro categoría de conductas negativas hacia el docente</i>	
NPA	No presta atención en las explicaciones del docente.
NAS	No colabora con el docente en la organización de la tarea.
DIS	Lleva a cabo conductas disruptivas con el docente.
EFP	Lleva a cabo conductas de provocación con el docente.
NPAR	No desarrolla las tareas de forma inmediata, se inhibe en las mismas.
<b>Criterio 3. Conductas hacia el equipamiento y el material</b>	
<i>3.1. Macro categoría de conductas positivas hacia el material/instalaciones</i>	
CI	Conserva y conoce las instalaciones donde se desarrolla la tarea.
CVRG	Utiliza el material de forma debida, lo ubica de forma adecuada.
<i>3.2. Macro categoría de conductas negativas hacia el material/instalaciones</i>	
NCI	No conserva ni conoce las instalaciones donde se desarrolla la tarea.
NCVRG	No utiliza el material de forma debida, no lo ubica adecuadamente.
<b>Criterio 4. Conductas hacia las reglas-normativa</b>	
<i>4.1. Macro categoría de conductas positivas hacia las reglas/normativa</i>	
RCR	Acepta y reconoce el reglamento del deporte/actividad a desarrollar.
PT	Para la tarea para auxiliar, asistir a un participante lesionado.
RT	Retoma la tarea en la situación en la que se detuvo.
IT	Identifica y lleva a cabo acciones de estrategia, fundamentales en la tarea.
<i>4.2. Macro categoría de conductas negativas hacia las reglas/normativa</i>	
NRCR	No acepta el reglamento del deporte/actividad a desarrollar.
NPT	No para la tarea para auxiliar, asistir a un participante lesionado.
NRIT	No retoma la tarea en la situación en la que se detuvo, para obtener ventaja.
JP	Realiza acciones peligrosas con su conducta.
PTI	Deja pasar el tiempo deliberadamente.
NIT	No lleva a cabo acciones de estrategia, fundamentales en la tarea.
<b>Criterio 5. Faltas por contacto.</b>	
FG	Conducta antirreglamentaria por juego duro.
FL	Conducta antirreglamentaria por juego brusco.

Tabla 1. Criterios, Macro categorías y Categorías de la Herramienta de Observación.

un contexto natural. El registro de los datos es de carácter continuo lo que implica que no existe ninguna fisura en la continuidad del flujo de conductual, de acuerdo con los planteamientos de Hernández Mendo (1996). En cuanto al grado de participación del observador podemos subrayar que el presente estudio toma como criterio la participación de los observadores. El procedimiento de registro fue la *codificación directa* a través de la hoja de calculo Excel XP, realizada por dos equipos de observación compuestos cada uno de ellos por tres observadores, entrenados previamente en un protocolo observacional. Se emplearon tres sesiones de treinta minutos de duración cada una, correspondientes a la parte central de la sesión de EF, no realizándose observaciones en el calentamiento y en la vuelta a la calma.

Cada uno de los grupos utilizó la concordancia consensuada (Anguera, 1990), en todas las ocasiones. Uno de los dos grupos de observadores realizó dos veces la misma observación, con un período intermedio de 15 días, con el objetivo de realizar posteriormente un análisis de la calidad del dato intraobservadores.

La inobservabilidad estimada es de tipo tecnológico (Hernández Mendo, 1996), y en ningún caso ha superado la ruptura de la continuidad de la sesión de observación por un periodo superior al 10% del total de la misma (Anguera, 1990), por lo cual no se ha estimado su influencia en la presente investigación. En cuanto a los requisitos de constancia *intra-intersesional*, se garantizó al máximo, escogiéndose sesiones donde no existían circunstancias excepcionales de

Definición del NÚCLEO CATEGORIAL	Grado de Apertura	Kohlberg	Tipo de Actividad
Integración/I: El participante observado lleva a cabo una conducta orientada a introducir en la actividad al resto de los participantes. Acepta a los demás participantes, como componentes de su grupo.	<p>2.1 Facilita la práctica de sus iguales, les presta el material de forma rápida y precisa o de motu propio. Les enseña las habilidades a desarrollar.</p> <p>2.2. Cuando un participante no comprende las normas de una tarea, le ayuda explicándosela por gestos o de forma verbal, e incluso cuando no se le solicita.</p> <p>2.3.No critica a otros participantes con gestos o acciones despectivas cuando fallan en la realización de una tarea, acepta la inclusión de otros participantes en el equipo, y cuando se forman los grupos, asienten con la cabeza o muestran su conformidad con gestos de aprobación. No realiza gestos de rechazo durante la realización de las tareas respecto a los participantes de su equipo con menor habilidad.</p> <p>2.4. Acepta su ubicación dentro de la tarea, acepta su posición y lugar dentro del grupo, llevando a cabo la tarea correspondiente. No realiza malos gestos, o expresiones de enfado ante su posición dentro del grupo.</p>	Estadio 5. Derechos y contrato social.	Colaboración. Colaboración-Oposición.

Tabla 2. Descripción completa de la Categoría I. (Introducción en la actividad al resto de participantes), correspondiente al criterio 1, conductas hacia sus iguales, del Macrosistema, conductas positivas hacia sus iguales.

ruptura del flujo conductual, tales como caída de postes, apagones de luz, accidentes o ruptura del material auxiliar etc. Y en cuanto a la constancia intersesional, además de ser las sesiones propias del contenido didáctico de deportes colectivos, el Voleibol, en su desarrollo se han seguido los objetivos, contenidos, metodologías propias del contexto educativo.

## Resultados

Se han estimado las correlaciones con los coeficientes de *Pearson*, *Tau* de *Kendall* y de *Spearman*. Estos índices se han determinado a partir de una tabla de frecuencias de todas las categorías objeto de estudio, obtenidas por cada equipo de observadores. Las observaciones 1 y 2, corresponden a las realizadas por el primer grupo de observadores en dos momentos diferentes, esto permite estimar las concordancias *intra-observadores*. La observación 3, corresponden al segundo grupo de observadores, esto permite estimar las concordancias *inter-observadores*. (ver Tabla 3).

Los resultados del análisis de correlaciones han sido satisfactorios y se sitúan entre el 0.98 y 1.00. La *concordancia intra-observadores* obtuvo los siguientes resultados: coeficiente de *Pearson*= 1.00, *Tau* de *Kendall*=0.98, *Spearman*= 0.99. La *concordancia inter-observadores*: coeficiente de *Pearson*=1.00, *Tau* de *Kendall*=0.98, *Spearman*=0.99.

El índice *kappa* de *Cohen* (1960) se utiliza cuando los datos son categóricos, o cuando la variable de respuesta se clasifica de acuerdo con una escala nominal o multinomial, una medida de acuerdo entre observadores, similar a la correlación *intra-clase*. El índice *kappa* es un estadístico de concordancia que corrige el azar. Este coeficiente se ha utilizado para estimar la fiabilidad por criterios y el global de la sesión.

Los resultados arrojan unos valores de 0.91 para la *concordancia intra-observadores*, y de 1,00 para la *concordancia inter-observadores*. *Fleiss* (1981) caracteriza como regulares los valores de *kappa* que se hallan entre 0.40 y 0.60, buenos de 0.60 a 0.75, y excelentes por encima de 0.75. Como complemento a esto, se estimo al alfa *Cronbach*, que calculado como competencia de observadores obtuvo un resultado de 1 (Losada y Arnau, 2000).

La *Teoría de la Generalizabilidad* pretende comprobar que las puntuaciones obtenidas sean representativas de un conjunto más amplio. Siendo los objetivos para los que se ha hecho uso de esta teoría los siguientes: a) Determinar la fiabilidad de los observadores. b) Valorar la homogeneidad de las categorías o del sistema de categorías. c) Optimizar el diseño de medida, esto nos permitirá conocer el número mínimo de sesiones necesarias para generalizar con precisión cualquier resultado procedente de la investigación que tuviera en cuenta las características definidas en la muestra (Blanco, 1997; Blanco Villaseñor y Anguera, 1991; Blanco Villaseñor y Hernández Mendo, 1998)

Tal y como aparecen en la tabla número 4, para el cálculo de la *fiabilidad inter-observadores* se empleó un diseño de dos facetas (categorías/observadores = C/O). La determinación de las fuentes de varianza reveló que la mayor parte de la variabilidad (100%), quedaba asociada a la faceta categorías, siendo nula en el resto de facetas, *observadores*, y *observadores/categorías* (0%). El análisis de los coeficientes de generalizabilidad arrojaron resultados que indicaban una alta fiabilidad en la precisión de generalización de los resultados (1.00). Tal y como aparecen en la tabla número 4, para

el cálculo de la *homogeneidad de las categorías* se empleó un diseño de dos facetas (observadores/categorías = O/C). La determinación

de las fuentes de varianza fue la misma que para el diseño anterior. El resultado de los coeficientes de fiabilidad (0.00) y de generali-

Frecuencias de conductas.			
CATEGORÍAS	OBSERVACIÓN 1	OBSERVACIÓN 2	OBSERVACIÓN 3
A	0	0	0
I	118	116	118
CO	108	108	111
S	60	55	57
AP	34	34	36
DM	123	123	123
R	119	119	119
NDM	7	7	7
AV	5	5	5
NI	1	2	1
NCO	4	4	4
IN	8	8	8
NR	12	11	11
PA	58	59	59
AS	124	123	125
PAR	113	114	114
NPA	6	6	6
NAS	4	5	4
DIS	8	10	8
EFP	0	0	0
NPAR	7	7	7
CI	129	129	129
CVRG	125	125	125
NCI	1	1	1
NCVRG	1	1	1
RCR	121	121	121
PT	3	3	3
RT	110	110	110
IT	49	49	49
NRCR	10	10	10
NPT	1	1	1
NRIT	0	1	0
JP	6	6	6
PTI	5	5	5
NIT	0	0	0
FG	0	0	0
FL	3	3	3
TOTALES	1483	1481	1487

Tabla 3. Frecuencias correspondientes a los distintos momentos de observación para el análisis de calidad del dato.



zabilidad (0.00) indican que no existe superposición y que se evalúa todo el espacio categorial, y por tanto, que el sistema de categorías es exhaustivo y mutuamente excluyente. Asimismo, en la tabla número 4, los índices de generalizabilidad y de fiabilidad, para una optimización del diseño de medida, se observa que con 3 sesiones, se

obtienen índices de 0.978, en ambos. Si se consideran 6 sesiones, los índices suben hasta 0.989 y con 9 sesiones, los índices se sitúan en 0.993. Con lo que podemos aseverar, que para 3 sesiones, presenta unos índices de generalizabilidad y fiabilidad óptimos, los cuales permiten asegurar la fiabilidad, precisión y validez de la herramienta.

<b>Fiabilidad inter-observadores. Diseño C/O.</b>					
	Suma	G.L.	Media	Error	%
	Cuadrados		Cuadrados	Estandar	
O	0.55	2	273	0.00557	0
C	281231.10	36	7811975	697.39666	100
CO	31.47	72	437	0.07185	0
I. Fiabilidad			1		
I. Generalizabilidad			1		
<b>Homogeneidad de las categorías. Diseño O/C.</b>					
O	0.55	2	0.273	0.00557	0
C	2.81231.10	36	784.975	697.39666	100
OC	31.47	72	0.43707	0.07185	0
I. Fiabilidad			0		
I. Generalizabilidad			0		
<b>Optimización del Diseños de Medida. Diseño C/S.</b>					
	Niveles	Niveles Procesados	Tamaño	Tamaño Muestra de Estudio	Tamaño Muestra de Estudio
S	3	3	Infinito	6	9
C	37	37	Infinito	37	37
Total número de Observaciones		111		222	333
I. Fiabilidad				0.989	0.993
I. Generalizabilidad				0.989	0.993

Tabla 4. Resultados Análisis de generalizabilidad.

## Discusión

La importancia de este trabajo, se centra en la construcción y validación de una herramienta "ad hoc", la cual permita obtener información precisa, válida y fiable para posteriores estudios científicos en el ámbito del desarrollo positivo, en las clases de EF de la ESO. La literatura científica, apenas proporciona estudios relevantes en el ámbito de la metodología observacional al respecto. La base

de datos *Pro-Quest*, en múltiples bases de datos, y con un filtro exclusivo acerca de revistas científicas, realizando la búsqueda en los términos (*physical education and observation, physical education and observation and moral development*) obtiene 53 registros, de los cuales ninguno se sitúa en la temática desarrollada en el presente estudio, proporcionando referencias de artículos sobre salud, entrenamiento, recreación, etc. En cuanto al segundo criterio de búsqueda, no arroja ningún resultado.

Respecto a otras herramientas de observación utilizadas en diferentes estudios y trabajos hay que señalar que presentan dificultades en cuanto a la estimación de la calidad del dato se refiere. La herramienta de observación, utilizada en el programa de Hellinson, (1995) a través de el programa TPRS. (*Tomando responsabilidad personal y social*), de Escartí, Pascual y Gutiérrez, (2005), presenta exclusivamente una concordancia consensuada, no profundizando en los análisis de calidad del dato.

El "Instrumento de Observación de conductas de Fair Play en Fútbol", (IOFF) de Cruz, Boixadós, Valiente y Torregrosa (1999), presenta un índice Kappa de Cohen de 0,57. Los resultados aquí presentados obtienen mejores indicadores.

En cuanto a la herramienta presentada por Planchuelo, Fernández García y Hernández Mendo (2007), el abordaje realizado guarda características comunes en el análisis de calidad del dato (Planchuelo, Hernández Mendo y Fernández García, 2009). Esta similitud abarca desde la concordancia consensuada (Anguera, 1990), el entrenamiento de los observadores, la elaboración de archivos de detección de errores, (usando programa de SDIS-GSEQ de Bakeman y Quera, 1996), la determinación de coeficientes de concordancia intra-observadores e interobservadores, hasta los estudios de generalizabilidad de los datos (Blanco y Hernández, 1998).

Se puede afirmar que la herramienta observacional presentada, obtiene en cuanto a la calidad del dato, unos resultados satisfactorios, tanto en los coeficientes de correlación, como los de asociación y en los índices de fiabilidad y generalizabilidad. Así pues verificamos que el sistema mixto propuesto, cumple con los requisitos metodológicos relativos a la fiabilidad, precisión y

validez exigibles. Podemos así, obtener información de los diferentes aspectos de las conductas de las clases de EF, conductas relacionadas con el desarrollo positivo.

Respecto a la posterior aplicabilidad de dicha herramienta, consideramos que requiere de una complementariedad metodológica. Existe un amplio debate entre las metodologías tradicionalmente denominadas cualitativa y cuantitativa. De forma alternativa, los defensores de cada uno de ambos polos se empeñaban en radicalizar las críticas a la posición alternativa. En las últimas décadas ha ido evidenciándose la reducción del enfrentamiento, y se han abierto vías para lograr la complementariedad, a la vez que se abre un interrogante respecto a la integración. La metodología observacional, por sus peculiares características, constituye un vivo ejemplo de dicha complementariedad entre lo cualitativo y lo cuantitativo (Anguera, 2004); no obstante la propuesta que aquí se plantea es la relativa no solo a la utilización de procedimientos de distintas metodologías sino también al uso de la triangulación ya sea de datos, del investigador, de la teoría o de la metodología (inter o intra métodos) (Denzin, 1989; Flick, 2007). Se hace necesario pues, complementar otros estudios realizados en el ámbito del desarrollo moral a partir de esta consideración (Cechini, Montero y Peña, 2003; Dias, Cruz y Danish; 2000; Vivó, 2001); y planificar programas con sus correspondientes evaluaciones que recojan esta propuesta (Planchuelo, 2008; Planchuelo, Hernández Mendo y Fernández García, 2009). Consideramos viable realizar este tipo de evaluación de programas aplicados en el ámbito de la docencia de la EF y su incidencia en el desarrollo moral de los participantes en función de diversos criterios (p. e. objetivos competitivos frente a objetivos cooperativos).

*CONSTRUCCIÓN DE UNA HERRAMIENTA OBSERVACIONAL PARA EVALUAR LAS CONDUCTAS PROSOCIALES EN LAS CLASES DE EDUCACIÓN FÍSICA*

**PALABRAS CLAVE:** Conducta prosocial, Observación, Educación Física.

**RESUMEN:** El presente trabajo describe la construcción y validación de un sistema taxonómico original, "ad hoc", que permite observar las conductas prosociales que se producen en las clases de Educación Física (EF) en la Enseñanza Secundaria Obligatoria (ESO). Dichas conductas se encuentran relacionadas con el desarrollo de valores positivos. Técnicamente, es una herramienta que combina los sistemas de formatos de campo y los sistemas de categorías E/ME. Este sistema se ha desarrollado desde una perspectiva deductiva desde el punto de vista del desarrollo del razonamiento moral de Kohlberg. El proceso de optimización de la herramienta, incluye el control y análisis de calidad del dato, destacando la concordancia consensuada y el análisis de la teoría de generalizabilidad, los cuales arrojan unos resultados óptimos.

*CONSTRUÇÃO DE UMA FERRAMENTA OBSERVACIONAL PARA AVALIAR CONDUTAS PROSOCIAIS NAS AULAS DE EDUCAÇÃO FÍSICA*

**PALAVRAS-CHAVE:** Comportamentos pró-sociais, Observação, Educação Física.

**RESUMO:** O presente trabalho descreve a construção e validação de um sistema taxonómico original, "ad hoc", que permite observar comportamentos pro-sociais que se produzem nas aulas de Educação Física (EF) no Ensino Secundário Obrigatório (ESO). Estes comportamentos encontram-se relacionados com o desenvolvimento de valores positivos. Tecnicamente, é uma ferramenta que combina os sistemas de formato de campo e os sistemas de categorias E/ME. Este sistema foi desenvolvido a partir de uma perspectiva dedutiva baseada no desenvolvimento moral de Kohlberg. O processo de optimização da ferramenta inclui o controlo e análise da qualidade de dados, destacando-se a consistência e a análise consensual da teoria da generalização, os quais apontam óptimos resultados.

## Referencias

- Anguera, M. T., Hernandez Mendo, A., Gorospe Egaña, G. y Martínez de Santos, R. (2005). Desarrollo y optimización de una herramienta observacional en el tenis de individuales. *Psicothema*, 17 (1), 123-127.
- Anguera, M. T. (2004). Posición de la metodología observacional en el debate entre las opciones metodológicas cualitativa y cuantitativa. ¿Enfrentamiento, complementariedad, integración? *Psicología em Revista (Belo Horizonte, Brasil)*, 10 (15), 13-27.
- Anguera, M. T. (1979). Observación de la conducta espacial. Comunicación presentada al *VI Congreso Nacional de Psicología. Pamplona*.
- Anguera, M. T. (1990). Metodología observacional. En J. Arnau, M. T. Anguera y J. Gómez, *Metodología de la investigación en Ciencias del Comportamiento* (pp. 125-236). Murcia: Secretariado de Publicaciones de la Universidad de Murcia.
- Anguera, M. T. (1993). Proceso de Categorización. En Anguera Argilaga, M. T. (Ed.): *Metodología Observacional en la investigación psicológica. Vol I: Fundamentación*. Barcelona: PPU.

- Anguera, M. T., Blanco Villaseñor, A. y Losada López, J. L. (2001). Diseños observacionales, cuestión clave en el proceso de la metodología experimental. *Metodología de las Ciencias del Comportamiento*, 3 (2) 135-160.
- Anguera, M. T., Blanco Villaseñor, A., Losada López, J. L. y Sánchez Algarra, P. (1999). Análisis de la competencia en la selección de observadores. *Revista Metodología de las Ciencias del Comportamiento*, 1, 95-114.
- Anguera, M. T. y Blanco Villaseñor, A. (2003). Registro y codificación del comportamiento deportivo. En Hernández Mendo, A. (Ed.), *Psicología del Deporte (vol II): Metodología* (pp. 6-349). Buenos Aires: Tulio Guterman ([www.efdeportes.com](http://www.efdeportes.com)).
- 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, agosto 2000. <http://www.efdeportes.com/efd24b/obs.htm> [Consulta: 10 de octubre de 2000]
- Ardá Suárez, A. (1998). *Análisis de los patrones de juego en fútbol a 7. Estudio de las acciones ofensivas*. Universidade da Coruña: Tesis Doctoral no publicada.
- Bakeman, R. y Quera, V. (1996). *Análisis de la interacción. Análisis secuencial con SDIS y GSEQ*. Madrid: Rama.
- Blanco Villaseñor, A. (1989). Fiabilidad y generalización de la observación conductual. *Anuario de Psicología*, 43 (4), 5-32.
- Blanco Villaseñor, A. (1993). Fiabilidad, precisión, validez y generalizabilidad de los diseños observacionales. En Anguera Argilaga M.T. (Ed.), *Metodología observacional en la investigación psicológica: Vol. 2: Fundamentación (2)* (pp. 151-261). Barcelona: PPU.
- Blanco Villaseñor, A. (1997). Precisión en la evaluación de la Investigación Observacional. Comunicación presentada al V Congreso de Metodología de las CC. Humanas y Sociales. Sevilla: AEMCCO.
- Blanco Villaseñor, A. y Anguera Argilaga, M. T. (1991). Sistemas de codificación. En Anguera Argilaga, M. T. (Ed.), *Metodología observacional en la investigación psicológica* (pp. 193-239). Barcelona: P.P.U.
- Blanco Villaseñor, A. y Hernández Mendo, A. (1998). Estimación y generalización en un diseño de estructura espacial defensiva en el fútbol. En J. Sabucedo, R. García Mira, E. Ares y D. Prada, *Medio Ambiente y Responsabilidad Humana* (pp. 579-583). A Coruña: Tórculo.
- Blanco Villaseñor, A., Castellano Paulis, J. y Hernández Mendo, A. (2000). Generalización de las observaciones de la acción del juego en el fútbol. *Psicothema*, 12 (2 supl), 81-86.
- Boixadós i Anglés, M. (1998). *Evaluación del clima motivacional y de las actitudes de Fairplay en futbolistas alevines e infantiles y los efectos de el asesoramiento psicológicos, a los entrenadores*. Tesis Doctoral. Universidad Autónoma de Barcelona.
- Bredemeier, B. (1991). Morality and sport for all. En P. Oja, y R. Telama (Eds.), *Sport for all* (pp. 365-372). Ámsterdam: Elsevier Science Publishers,.
- Castellano Paulis, J. (2000). *Observación y análisis de la acción del juego en el fútbol*. Universidad del País Vasco: Tesis Doctoral.
- Castellano Paulis, 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), pp. 635-641.

- Castellano Paulis, J., Hernández Mendo, A., Haro Romero, J. A. (2002). Mapas socioconductuales de la selección francesa en el mundial de fútbol de Francia' 98. *Revista de Psicología del Deporte*, 11 (1), 35-51.
- Cecchini Estrada, J. A., Montero Fernández, J. y Peña Calvo, J. V. (2003). Repercusiones del Programa de Intervención para Desarrollar la Responsabilidad Personal y Social de Hellinson sobre los comportamientos de Fairplay y el auto-control. *Psicothema*, 15 (4), 631-637.
- Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 20, 37-46.
- Cruz i Feliú, J., Boixadós i Anglés, M., Valiente i Barros, L. y Torregrosa, M. (1999). ¿Se pierde el fairplay y la deportividad en el deporte en edad escolar?. *Apunts. Educación Física y Deportes*, 64, 6-16.
- Denzin, N. K. (1989). *The research Act*. Englewood Cliffs, NJ: Prentice Hall.
- Días, C., Cruz, J. y Danish, S. (2000). El deporte como contexto para el aprendizaje y la enseñanza de competencias personales. Programas de intervención para niños y adolescentes. *Revista de Psicología del Deporte*, 9 (1-2), 107-122.
- Escartí Carbonell, A. y García Martín, A. (1994) Factores de los iguales relacionados con la practica y la motivación deportiva en la adolescencia. *Revista de Psicología del Deporte*, 3 (6), 35-53
- Escartí Carbonell, A., Pascual Baños, C. Y Gutiérrez Sanmartín, M. (2005). *Responsabilidad personal y social a través de la educación física y el deporte*. Barcelona: Grao.
- Fleiss, J. L. (1981). *Statistical methods for rates and proportions*. Nueva York: Wiley.
- Flick, U. (2007). *Introducción a la investigación cualitativa*. Madrid: Morata.
- Garay Plaza, J. O., Hernández Mendo, A. y Morales Sánchez, 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.
- Garay Plaza, J. O., Hernández Mendo, A. y Morales Sánchez, V. (2007). Análisis secuencial en el tenis de dobles. *Revista de Psicología General y Aplicada*, 60 (3), 253-269.
- Gibbons, S., Ebbeck, V. y Maureen, R. (1995). Fair Play for Kids: Effects on the moral development of children in physical education. *Research Quarterly for Exercise and Sport*, 66 (3), 247-255.
- Gutiérrez Sanmartín, M. y Escartí Carbonell, A. (2006). Influencia de padres y profesores sobre las orientaciones de meta de los adolescentes y su motivación intrínseca en educación física. *Revista de Psicología del Deporte*, 15 (1), 23-35.
- Hellinson, D. R. (1995). *Teaching personal and social responsibility through physical activity*. Champaign, IL: Human Kinetics.
- Hernández Mendo, A. (1996). *Observación y análisis de patrones de juego en deportes sociomotores*. Universidad de Santiago de Compostela: Tesis Doctoral.
- Hernández Mendo, A. (2000). La metodología observacional en el voleibol. *Lecturas: EF y Deportes. Revista Digital*, 25, septiembre. <http://www.efdeportes.com/efd25b/obsvol.htm> [Consulta: 20 de octubre de 2000].
- Hernández Mendo, A. y Anguera Arguilaga, M. T. (1997). Aportaciones del análisis secuencial a las acciones de juego en deportes sociomotores. Comunicación presentada al V Congreso de Metodología de las CC. Humanas y Sociales. Sevilla: AEMCCO.

- Hernández Mendo, A., Aragundi, C. A. y González Fernández, M. D. (1995b). Análisis de calidad de datos en registros observacionales en voleibol. En M. T. Vega y M. C. Taberero, *Psicología Social de la educación y de la Cultura, Ocio, Deporte y Turismo* (pp. 423-427). Salamanca: Eudema.
- Hernández Mendo, A., Areces, A., Vales, A. y González Fernández, M. D. (1995a). Análisis de calidad del dato en registros observacionales de deportes sociomotores: fútbol. En M. Ato y J. A. López Pina (Eds.). *Comunicación presentada al IV Symposium de Metodología de las Ciencias del Comportamiento* (pp. 143-147). Murcia: Servicio de Publicaciones de la Universidad de Murcia.
- Hernández Mendo, A., Areces, A., Vales, A. y González Fernández, M. D. y Garea, J. (1994). Observación conductual en el hockey sobre patines. Comunicación presentada al *I Congreso Internacional de Psicología Conductual. I International Congress of Behavioral Psychology*.
- Kohlberg, L. (1981). *Essays on Moral Development. vol. 1: The Philosophy of Moral Development*. San Francisco CA: Harper & Row.
- Losada, J. L. y Arnau, J. (2000). Fiabilidad entre observadores con datos categóricos mediante el anova. *Psicothema, 12 (Supl. 2)*, 335-339.
- Lee, M. (1988). Values and Responsibilities in Children's Sports. *Physical Education Review, 11*, 19-27.
- Marías, J. (1941). *Historia de la filosofía*. Madrid: Alianza Editorial.
- Planchuelo, L. (2008). *Intervención y evaluación del desarrollo moral en educación física en primaria*. Tesis Doctoral: Universidad de Málaga.
- Planchuelo Martín, L., Fernández García, J. C. y Hernández Mendo, A. (2007). Herramienta Observacional para evaluar el desarrollo moral de los niños en las clases de EF. *Symposium de observación en la actividad física y deporte. Actas del X Congreso de Metodología de las Ciencias Sociales y de la Salud*. Barcelona.
- Planchuelo Medina, L., Hernández Mendo, A. y Fernández García, J. C. (2009). Intervención y evaluación de un Programa de Desarrollo Moral en la Educación Física en Primaria. . *Lecturas: EF y Deportes. Revista Digital, 14 (133), junio*. <http://www.efdeportes.com/efd133/programa-de-desarrollo-moral-en-la-educacion-fisica.htm> [Consulta: 22 de diciembre de 2008].
- Romance, T. J. (1984). *A Program to promote Moral Development Through Elementary School Physical Education*. Tesis Doctoral. Oregon University.
- Vivó, P. (2001). *Desarrollo de valores y razonamiento moral a través de la educación física en educación secundaria*. Tesis Doctoral: Universidad de Valencia.
- Weiss, M. R. y Smith, A. L. (1999). Quality of youth sport friendships: Measurement development and validation. *Journal of Sport and Exercise Psychology, 21*, 145- 166.
- Ysewijn, P. (1996). *GT software for generalizability studies*. Mimeografía.