

UNA HERRAMIENTA OBSERVACIONAL PARA LA EVALUACIÓN DEL DESARROLLO MORAL EN LAS CLASES DE EDUCACIÓN FÍSICA EN PRIMARIA

Antonio Hernández Mendo¹ y Lina Planchuelo Medina²
Universidad de Málaga¹ y CEIP Cristo de Mena. Consejería de
Educación. Junta de Andalucía²

RESUMEN: La herramienta observacional que presentamos ha sido usada en la observación de un programa de actividad física que tiene como objetivo la mejora del crecimiento moral del alumnado en Primaria. La muestra fue recogida en un centro de la provincia de Málaga. Está formada por alumnas/os de niveles de 3º a 5º de Primaria. El diseño observacional utilizado es Seguimiento/Nomotético/Multidimensional (S/N/M), en el cuadrante IV. En este trabajo se presenta el proceso de elaboración y optimización de la herramienta "ad hoc" que está formada por un sistema mixto de formatos de campo y sistema de categorías. En este trabajo se aborda una doble problemática, la utilización de la Metodología Observacional (MO) y el uso de Teoría de la Generalizabilidad (TG) en la evaluación del desarrollo moral en las clases de educación física.

PALABRAS CLAVE: desarrollo moral, observación, Educación Física, Educación Primaria.

OBSERVATIONAL TOOL FOR ASSESSMENT OF MORAL DEVELOPMENT IN THE PHYSICAL EDUCATION CLASSES IN PRIMARY

ABSTRACT: The present observational tool has been used in the observation of a physical activity program that aims to improve the moral growth of students at primary level. The sample was collected in a center of the province of Malaga, is formed by students in levels 3 through 5th grade. The observational design used is Monitoring / Nomothetic / Multidimensional (S / N / M), in quadrant IV. This paper presents the process of elaboration and optimization tool "ad hoc" which consists of a mixed format field and category system. This paper addresses two problems: the use of the Observational Method (OM) and the use of Generalizability Theory (GT) in the assessment of moral development in physical education classes.

KEYWORDS: moral development, observation, Physical Education, Primary Education.

UMA FERRAMENTA OBSERVACIONAL PARA AVALIAR O DESENVOLVIMENTO MORAL NA EDUCAÇÃO FÍSICA NA PRIMÁRIA

RESUMO: O presente ferramenta de observação foi utilizada na observação de um programa de atividade física que visa melhorar o crescimento moral dos alunos na Primária. A amostra foi coletada em um centro da província de Málaga. É composto por estudantes dos níveis mínimos 3 a 5^a série. O desenho observacional utilizado é Faixa / nomotética / Multidimensional (S / N / M), no quadrante IV. Neste trabalho, a ferramenta de preparação e otimização "ad hoc", que consiste de um sistema misto de formatos de campo e sistema de categoria. Este artigo aborda dois problemas utilizando metodologia observacional (MO) e do uso da teoria da generalização (TG) na avaliação do desenvolvimento moral em aulas de educação física.

PALAVRAS-CHAVE: desenvolvimento moral, a observação, Educação Física, Ensino Fundamental.

El desarrollo moral de los niños, es crucial para la sociedad actual y para el sistema educativo. Este último tiene como función primordial la educación integral de las personas y la formación en valores (y por tanto en desarrollo moral). Las sociedades occidentales están acusando la crisis de valores a las que están sometidas, y han apostado por la búsqueda de cambios eficaces y efectivos en los programas curriculares educativos. Estos cambios, han ido encaminados a conceder a la educación moral, mayor protagonismo. En las últimas décadas, han aumentado el número de investigaciones científicas en éste ámbito, y más concretamente en el área de educación física (EF). Éstas han demostrado la importancia de la EF, y el deporte, como un proceso óptimo para promover el desarrollo moral en los niños y adolescentes (Bredemeier, 1991; Gibbons, Ebbeck y Maureen, 1995; Hellinson, 1995; Lee, 1988; Romance, 1984; Weiss y Smith, 1999; Escartí y García, 1994; Gutiérrez y Escartí, 2006).

Este trabajo forma parte de una investigación más amplia cuyo objetivo era determinar si un Programa de Educación Física, diseñado para tal fin, incrementaba el desarrollo moral del alumnado de Primaria. Además, se aborda una doble problemática, por una lado, la utilización de la Metodología Observacional (MO), y, por otro, el uso de Teoría de la Generalizabilidad (TG) (Hernández Mendo, Díaz Martínez y Morales, 2010).

El uso de la MO para la evaluación del desarrollo moral en las clases de educación física tiene alcances epistemológicos importantes referidos al objeto de medida, y por tanto, a las implicaciones teóricas en la construcción de su herramienta, así como su consideración como técnica y no como método. Por tanto es importante considerar los siguientes aspectos (Anguera y Blanco, 2006):

(1) Categorización y codificación, son procesos interdependientes, la categorización responde al principio fundamental de reducción de la complejidad, y la codificación es la operación metodológica consistente en asignar símbolos o códigos a *unidades* de información procedentes de la realidad estudiada, que también cumple con el mismo principio.

(2) El registro, es imprescindible que tenga la propiedad transitiva, esto es, un registro codificado deberá poder decodificarse, transformándose en un registro no codificado sin que se produzca pérdida o deformación de la información.

(3) En MO existen dos herramientas fundamentales: el Sistema de Categorías E/ME (SC) y los Sistemas de Formatos de Campo (SFC).

(a) El sistema de categorías constituye la base inicial en el proceso de construcción del instrumento y consiste en un registro sistematizado tipo lista de rasgos que debería idealmente cumplir una garantía de *presunción de exhaustividad* superando una *prueba de cautela* (durante no menos de 3 sesiones consecutivas no debe aparecer ninguna conducta que no este en el repertorio, en caso contrario debe iniciarse de nuevo el proceso).

(b) Formato de campo, a partir de cada uno de sus criterios se elabora un *catálogo*, que es igualmente un registro tipo lista de rasgos, caracterizado por no ser exhaustiva, pudiéndose añadir nuevas conductas de forma indefinida. Esta característica permite superar las críticas realizadas a los sistemas de categorías en los estudios de carácter longitudinal, en los cuales se produce una evolución de los procesos (evolución del lenguaje, prosocialidad, juego espontáneo con objetos, etc.), dado que la aparición progresiva de nuevas conductas haría inviable la obtención inicial de una lista cerrada.

(4) En la actualidad, cada vez cobra mayor relevancia la combinación de ambos instrumentos, de forma que la estructura básica sea de formato de campo, y cada uno de los criterios este compuesto por un sistema de categorías (elaborado si existe marco teórico y si se pueden llegar a cumplir las condiciones de exhaustividad y mutua exclusividad). El marco teórico confiere una indudable consistencia y estabilidad al sistema de categorías, que desde este punto de vista sería un instrumento más fuerte (Anguera, 1990; Garay, Hernández-Mendo y Morales, 2006, 2007).

(5) La observación como método tiene entidad científica para la obtención de conocimiento, a diferencia de su uso como técnica, que constituye únicamente una estrategia subordinada a las directrices de otra metodología (selectiva o experimental). En estudios anteriores, (Boixadós, 1998, Cruz, Boixadós, Valiente y Torregrosa, 1999; Escartí, Pascual y Gutiérrez; 2005) se ha usado la observación como técnica. Esta cuestión es fundamental para delimitar la toma de decisiones metodológicas como es el tipo de diseño observacional (Anguera, Blanco, Losada y Hernández-Mendo, 2000; Anguera, Blanco, Hernández-Mendo y Losada, 2011), el tipo de datos (Bakeman y Quera, 1996), etc. La elección del tipo de diseño es trascendental pues condicionará el tipo de dato, el muestreo y el tipo de análisis que se realizará (Anguera, Blanco, Hernández-Mendo y Losada, 2011). 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. El principal motivo se encuentra en la superficialidad de obtención de los datos, y, por tanto, en su carácter inconsistente. Los principales diseños observacionales son de disposición no estándar -coherentes con la flexibilidad de la metodología observacional-, y considerando la mejora en el proceso de sistematización y optimización de los datos, han surgido nuevas propuestas en el análisis de datos. Se ha establecido un planteamiento donde se cruzan las *unidades de estudio*, ideográfico (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, 1993; Anguera, Blanco y Losada, 2011; Anguera, Blanco, Hernández-Mendo y Losada, 2011).

En relación al segundo aspecto planteado, 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, Castellano y Hernández-Mendo, 2000). 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, 1993). El eje central de la TG se encuentra en los componentes de variancia, dado que su magnitud nos aporta información sobre las fuentes de error que están afectando una medición conductual (Blanco y Hernández-Mendo, 1998; Blanco, Castellano y Hernández-Mendo, 2000; Blanco, 1997).

En el presente trabajo, se expone el desarrollo y optimización de un instrumento, para el registro y codificación de las conductas prosociales que aparecen durante el transcurso de las sesiones de EF. 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 Piaget (1977) y Kohlberg (1981, 1992).

El desarrollo y optimización de la herramienta, incluye el control y análisis de calidad del dato, requisito imprescindible a fin de considerar la herramienta, como fiable y precisa (Anguera, 1979; Ardá, 1998; Castellano, 2000; Castellano, Hernández-Mendo y Haro, 2002; Gorospe, Hernández-Mendo, Anguera y Martínez de Santos, 2005; Hernández-Mendo, 1996). 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 variancia a partir del cual se realizó un análisis de generalizabilidad.

MÉTODO

Participantes

La muestra pertenece a un Centro de Educación Infantil y Primaria (C.E.I.P.) de una población rural de aproximadamente 3500 habitantes de la provincia de Málaga. El centro tiene dos líneas por nivel y una ratio que oscila entre los 19-28 niñas/os por aula. La muestra está constituida por alumnado distribuido en 6 grupos, de tercero a quinto (entre 8 y 11 años), de Educación Primaria. Todos los participantes pertenecen al mismo centro escolar. La muestra está formada por 47 niñas (47.9%), y 51 niños (52.1%). La distribución por género en la condición de control es de 25 niñas (48.1 %) y 27 niños (51.9 %). El grupo experimental 22 niñas (47.8%) y 24 niños (52.2 %). Los grupos control y experimental se asignaron al azar a partir de los grupos controles contruidos por el centro.

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, un

reproductor-grabador de DVD y un televisor. En el análisis de los datos se ha utilizado el paquete estadístico SPSS para Windows versión 14.0, *paquete estadístico SAS* (Schlotzhauer y Littell, 1997), la hoja de calculo *Excel XP* y el programa informático *Generalizability Study* versión 2.0.E. (Ysewijn, 1996), así como el programa informático *SDIS-GSEQ* (Bakeman y Quera, 1996), para Windows, software de digitalización, así como el material y las instalaciones utilizadas para el desarrollo de las sesiones de EF.

Procedimiento

La taxonomía conductual presentada es un sistema mixto de formatos de campo y sistemas de categorías E/ME (Anguera, 1979; Ardá, 1998; Castellano, 2000; Hernández-Mendo, 1996). Los formatos de campo garantizan el registro de varios aspectos de una misma realidad, y en cada criterio de este sistema se desarrolla un sistema de categorías (Anguera 1979), El sistema está constituido por 10 criterios y 39 categorías, cada una de ellas recoge el marco teórico proporcionado por los modelos de Piaget (1977) y Kohlberg (1981, 1992). Como puede apreciarse en las tabla 1 y 2.

Tabla 1

Sistema mixto de formatos de campo y sistema de categorías, su relación con los estadios psicológicos de Piaget y Kholberg, el tipo de actividad y el tipo de dato

Crite- rio	Núcleo categorial	Estadios psicológicos		Tipo de actividad
		Piaget	Kholberg	Tipo de datos
1. Conductas POSITIVAS dirigidas hacia sus iguales	1.1. Conducta de ayuda (A).	Codificación de las reglas	Convencional.	Colaboración-oposición Eventos multimodales
	1.2. Conducta de apoyo (Ap).	Codificación de las reglas	Convencional	Colaboración Eventos multimodales
	1.3. Acepta (Ac).	Cooperación naciente o codificación de las reglas	Convencional	Colaboración-Oposición Eventos multimodales
	1.4. Acepta puesto (Ao).	Codificación de las reglas	Convencional	Colaboración Eventos multimodales
	1.5. Apacible (Ab).	Codificación de las reglas	Convencional	Colaboración-Oposición Eventos multimodales

Criterio	Núcleo categorial	Estadios psicológicos		Tipo de actividad
		Piaget	Kohlberg	Tipo de datos
2. Conductas NEGATIVAS dirigidas hacia sus iguales	2.1. Agresión física (Af).	Egocentrismo	Pre-Convencional	Colaboración-Oposición
				Eventos multimodales
	2.2. Agresión verbal (Av) o amenaza.	Egocentrismo	Pre-Convencional	Colaboración-Oposición
				Eventos multimodales
	2.3. Protesta (Pr).	Egocentrismo o Cooperación naciente	Pre-Convencional	Colaboración-Oposición
				Eventos multimodales
	2.4. No comparte (Nc).	Egocentrismo o cooperación naciente	Pre-Convencional	Colaboración-Oposición
				Eventos multimodales
2.5. Falta de deportividad (Fd).	Egocentrismo o cooperación naciente	Pre-Convencional	Colaboración-Oposición	
			Eventos multimodales	
2.6. Incomoda (Ica).	Egocentrismo	Pre-Convencional	Colaboración-Oposición	
			Eventos multimodales	
2.7. Anima altercados (Aa).	Egocentrismo	Pre-convencional	Colaboración-Oposición	
			Eventos multimodales	
3. Conductas POSITIVAS dirigidas hacia el profesor.	3.1. Cooperar (Coo).	Cooperación naciente o codificación de las reglas	Convencional	Colaboración
				Eventos multimodales
	3.2. Activo (Ai).	Cooperación naciente o Codificación de las reglas	Post-convencional	Colaboración
				Eventos multimodales
3.3. Presta atención (Pa).	Cooperación naciente o codificación de las reglas	Convencional	Colaboración	
			Eventos multimodales	
3.4. Acepta sanciones (As).	Cooperación naciente o codificación de las reglas	Convencional o Post-Convencional	Colaboración	
				Eventos multimodales

Criterio	Núcleo categorial	Estadios psicológicos		Tipo de actividad
		Piaget	Kohlberg	Tipo de datos
4. Conductas NEGATIVAS dirigidas hacia el docente.	4.1. Desobediente (Db).	Egocentrismo	Pre-Convencional	Colaboración-Oposición
				Eventos multimodales
	4.2. Afronta (Afr).	Egocentrismo	Pre-Convencional	Colaboración-Oposición
				Eventos multimodales
	4.3. Discrepa más de una vez (Dr).	Egocentrismo	Pre-Convencional	Colaboración - Oposición
				Eventos multimodales
	4.4. Interrumpe (Ite).	Egocentrismo	Pre-Convencional	Colaboración-Oposición
				Eventos multimodales
5. Conductas POSITIVAS dirigidas hacia el equipamiento	5.1. Recoge (Rg).	Cooperación naciente	Convencional	Colaboración
				Eventos multimodales
	5.2. Conserva (Cv).	Cooperación naciente	Convencional	Colaboración
				Eventos multimodales
	5.3. Recupera (Rp).	Cooperación naciente	Convencional	Colaboración
				Eventos multimodales
6. Conductas NEGATIVAS dirigidas hacia el equipamiento	6.1. No recoge (Nrg).	Egocentrismo	Pre-Convencional	Colaboración-Oposición
				Eventos multimodales
	6.2. Deteriora material (Dm).	Egocentrismo	Pre-Convencional	Colaboración-Oposición
				Eventos multimodales
	6.3. No recupera (Nrp).	Egocentrismo	Pre-Convencional	Colaboración-Oposición
				Eventos multimodales
7. POSITIVA instalaciones	7.1. Cuida las instalaciones (Ci).	Cooperación naciente o codificación de las reglas	Convencional.	Colaboración
				Eventos multimodales

Criterio	Núcleo categorial	Estadios psicológicos		Tipo de actividad
		Piaget	Kholberg	Tipo de datos
8.C. NEGAT.-VA hacia instalación	8.1. Deteriora instalaciones (Di).	Egocentrismo	Pre-Convencional	Colaboración-Oposición
				Eventos multimodales
9. Conductas POSITIVAS dirigidas hacia las reglas del juego.	9.1. Participa (P).	Codificación de las reglas	Convencional.	Colaboración
	9.2. Deportividad (Dt).	Codificación de las reglas	Convencional	Eventos multimodales
				Colaboración
	9.3. Acata normas (An).	Codificación de las reglas	Post-Convencional	Colaboración
	9.4. Superación (S).	Cooperación naciente	Convencional	Eventos multimodales
Colaboración				
10. Conductas NEGATIVAS dirigidas hacia las reglas del juego.	10.1. Incumple (Im).	Egocentrismo	Pre-Convencional	Colaboración-Oposición
				Eventos multimodales
	10.2. Protesta (Pro).	Egocentrismo	Pre-Convencional	Colaboración-Oposición
				Eventos multimodales
	10.3. Inconstante (Inc).	Egocentrismo	Pre-Convencional	Colaboración-Oposición
				Eventos multimodales
	10.4. Impuntual (Imp).	Egocentrismo	Pre-Convencional	Colaboración-Oposición
				Eventos multimodales
	10.5. Falta de participación (Fp).	Egocentrismo	Pre-Convencional	Colaboración-Oposición
				Eventos multimodales
	10.6. Brusco (Br).	Egocentrismo	Pre-Convencional	Colaboración-Oposición
				Eventos multimodales

Tabla 2

Ejemplo de un criterio del sistema mixto de formatos de campo y sistema de categorías (con núcleo categorial y grado de apertura), su relación con los estadios psicológicos de Piaget y Kholberg, el tipo de actividad y el tipo de dato

Criterio	Núcleo categorial	Grado de apertura	Estadios psicológicos		Tipo de actividad
			Piaget	Kholberg	Tipo de datos
1. Conductas POSITIVAS dirigidas hacia sus iguales	<p>1.1. Conducta de ayuda (A). Los participantes observados llevan a cabo una conducta asistencial, auxilian e integran a los compañeros en clase.</p>	<ul style="list-style-type: none"> - El participante observado corre, gira o anda para auxiliar a un compañero lesionado, le consuela, le ayuda a levantarse, avisa al docente cuando se produce alguna caída o altercado. - Mira la zona del cuerpo lesionada. - Ayuda a los demás a integrarse en clase: les presta el material de motu propio, les enseña las habilidades, les explica como se hace, no los critica o los desprecia con gestos. - Si el participante observado ocasiona un daño a otro participante, se dirige hacia él, le presta ayuda, le acompaña a sentarse. - Si comete un fallo durante el desarrollo del juego pide disculpas mediante gestos. - Rectifica su conducta. - Avisa al compañero o compañera que le toca salir. - Cuando un compañero/a no comprende las normas o el funcionamiento de un juego, el participante observado le ayuda explicándoselo con gestos. - Le da el material en la mano o hace el intento de dárselo. - El participante observado se disculpa cuando ocasiona daño a un compañero/a en el desarrollo del juego. - No acusa a los demás. 	Codificación de las reglas	Convencional	Colaboración-oposición
					Eventos multimo- dales

Criterio	Núcleo categorial	Grado de apertura	Estadios psicológicos		Tipo de actividad
			Piaget	Kholberg	Tipo de datos
	<p>1.2. Conducta de apoyo (Ap). Los participantes observados animan a sus compañeros durante el desarrollo del juego.</p>	<ul style="list-style-type: none"> - El participante observado alienta a sus compañeros/as mediante gestos de aprobación y apoyo (coloca su mano en el hombro o en la espalda, lo recibe con un abrazo) -Si un compañero/a comete un fallo no realiza gestos de enfado o de reproche por el fallo cometido. - Acepta que los niños/as menos integrados entren en su equipo mediante gestos de aprobación o de amistad. - El participante observado anima a sus compañeros/as para mejorar el resultado, aplaude o jalea las acciones positivas de sus compañeros/as. - El participante observado actúa con deportividad hacia sus iguales, realiza gestos agradables y positivos hacia los demás: Palmadas, estrechar la mano, abrazos. - Le cambia el material si le es desfavorable. Salta de emoción cuando participan sus compañeros/as. - Realiza gestos positivos y agradables hacia otros participantes. 	Codificación de las reglas	Conven.	<p>Colaboración</p> <hr/> <p>Eventos multimediales</p>
	<p>1.3. Acepta (Ac). El participante observado admite las disculpas de un compañero o compañera.</p>	<ul style="list-style-type: none"> - Cuando un compañero/a se disculpa, el participante observado realiza gestos de asentimiento, no de enfado o disconformidad. 	Coope-ración naciente o codifi-cación de las reglas	Conven.	<p>Colaboración - Oposición</p> <hr/> <p>Eventos multimediales</p>

Criterio	Núcleo categorial	Grado de apertura	Estadios psicológicos		Tipo de actividad
			Piaget	Kholberg	Tipo de datos
	1.4. Acepta puesto (Ao). El participante observado acata el lugar que ocupa en el desarrollo del juego.	- El participante observado muestra conductas de aceptación del puesto que le toca sin expresar gestos de enfado o malos modos. - Le deja paso a los compañeros/as.	Codificación de las reglas	Conven.	Colaboración - Eventos multimodales
	1.5. Apacible (Ab). El participante observado evita participar en acciones violentas propias o de demás.	- El participante observado cuando hay un enfrentamiento, evita las acciones violentas no enfrentándose a los demás. - Si otros participantes se están enfrentando los separa. - Cuando un compañero/a se cuele en una fila no le pega, ni realiza gestos hoscos ni malhumorados. - Cuando otro participante le lanza la pelota lejos no llama la atención del docente ni se altera.	Codificación de las reglas	Conven.	Colaboración - Oposición - Eventos Multimodales

El tipo de datos utilizados son multieventos, describen las múltiples conductas que pueden sucederse en un contexto natural. El registro de los datos es continuo, sin fisuras en el flujo de conducta (Anguera y Blanco, 2003; Hernández-Mendo, 1996). La *codificación* se realizó a través de la hoja de cálculo Excel XP, realizada por dos equipos de observación compuestos cada uno de ellos por tres observadores, entrenados previamente con esta herramienta. 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. La codificación en cada uno de los grupos se realizó mediante concordancia consensuada (Anguera, 1990), en todas las ocasiones (forma cualitativa de control de calidad del dato).

La inobservabilidad descubierta es de tipo tecnológico (Hernández Mendo, 1996), en ningún caso ha superado el 10% del total (Anguera, 1990) y por tanto no ha

sido considerada. Se garantizó la constancia *intra-intersesional*, escogiendo sesiones donde no existían circunstancias excepcionales de ruptura del flujo conductual. 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

El análisis de calidad del dato se comenzó realizando un análisis de correlaciones con los coeficientes de Pearson, Tau de Kendall y de Spearman (Tabla 3). 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.. En primer lugar se procedió a realizar una codificación previa. Ésta se efectuó mediante la grabación de tres sesiones de observación, con el fin de contrastar las decisiones metodológicas. Dichas sesiones sirvieron para optimizar la herramienta observacional.

Tabla 3

Índices correlación

	E1-E2
Pearson	.999
Tau b de Kendall	.966
Rho de Spearman	.73

En segundo lugar se procedió a la realización del análisis Kappa de Cohen (1960), utilizando el programa estadístico SDIS-GSEQ (Bakeman y Quera, 1996), con el fin de estimar la fiabilidad interobservadores e intraobservadores. El cálculo se realizó en función de la distribución por grupos a fin de asegurar la homogeneidad en la fiabilidad de la herramienta con independencia del grupo. El índice *kappa* de Cohen (1960) se utiliza cuando los datos son categóricos, o cuando la variable de respuesta se clasifica en una escala nominal o multinomial, una medida de acuerdo entre observadores, similar a la correlación intraclase. El índice kappa es un estadístico de concordancia que corrige el azar. Los índices estimados para la *concordancia intra-observadores*, oscilan entre 0.93 y 0.80 (ver Tabla 4) y la *concordancia inter-observadores* oscilan entre 0.81 y 0.70. 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).

Tabla 4
Índices de Kappa de Cohen de los diferentes cursos

Cursos	Concordancia	Concordancia
	Intraobservadores	Interobservadores
3° A	.80	.81
4° A	.88	.81
5° B	.93	.70

En tercer lugar se ha llevado a cabo un estudio de generalizabilidad. Para determinar la fiabilidad inter e intra observadores se ha tomado un diseño de dos facetas (categorías y observadores). En el curso 3° A (que a posteriori, dentro del programa funcionaria como experimental) el análisis determina que el 98% de la variabilidad queda asociada a la faceta categorías. Siendo nula para la faceta observadores y de un 2% para la faceta de interacción categorías/observadores. El análisis del modelo C/O, revela un índice de Generalizabilidad excelente de un 0.995 (ver tabla 5). El modelo O/C presenta un buen ajuste.

Tabla 5
Análisis de generalizabilidad inter e intraobservadores 3° A

	Suma de cuadrados	GL	Media de cuadrados	%
O	0.05	2	0.026	0
C	4128.97	38	108.657	98
CO	43.28	76	0.570	2
		C/O	O/C	
I. Fiabilidad	0.995		0.000	
I. Generalizabilidad	0.995		0.000	
E. Relativo variancia	0.190		0.015	
E. Absoluto variancia	0.190		0.938	
Error estándar...	0.436		0.121	
Error estándar	0.436		0.969	

Para el análisis de Generalizabilidad de 4° A se toma el mismo diseño mencionado anteriormente. En este caso, el 99% de la variabilidad queda asociada a la variable categorías, es nula la variabilidad asociada a los observadores y de un 1% a la interacción de observadores categorías. El índice de Generalizabilidad es excelente, de un 0.996 (ver tabla 6).

Tabla 6

Análisis de generalizabilidad inter e intraobservadores 4° A

	Suma de cuadrados	GL	Media de cuadrados	%
O	0.02	2	0.009	0
C	5424.22	38	142.743	99
CO	44.65	76	0.587	1
	C/O		O/C	
I. Fiabilidad	0.996		0.000	
I. Generalizabilidad	0.996		0.000	
E. Relativo variancia	0.196		0.015	
E. Absoluto Variancia	0.196		1.230	
Error estándar ...	0.443		0.123	
Error estándar	0.443		1.109	

En este grupo experimental un 98% de la variabilidad queda asociada a las categorías, es nula para los observadores, siendo de un 2% a la interacción de categorías/observadores. En cuanto al análisis de los coeficientes de Generalizabilidad en esta estructura de diseño, determinan una fiabilidad de precisión de generalización de los resultados que consideramos bastante bueno, de un 0.995 (ver tabla 7).

Tabla 7

Análisis de generalizabilidad inter e intraobservadores 5° B

	Suma de cuadrados	GL	Media de cuadrados	%
O	3.35	2	1.675	0
C	8887.81	38	233.890	98
CO	95.32	76	1.254	2
	C/O		O/C	
I. Fiabilidad	0.995		0.251	
I. generalizabilidad	0.995		0.005	
E. Relativo Variancia	0.418		0.032	
E. Absoluto Variancia	0.422		2.020	
Error estándar...	0.647		0.179	
Error estándar	0.649		1.421	

DISCUSIÓN Y CONCLUSIÓN

La importancia de este trabajo, se centra en la construcción y validación de una herramienta “*ad hoc*”, que proporcione información precisa, válida y fiable y que ha permitido en un segundo estudio evaluar (en conjunto con otras herramientas) la eficacia y efectividad de un Programa de Desarrollo Moral en las clases de EF. En Primaria.

En diversos estudios se utilizan otras herramientas de observación, es preciso 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) denominado *Tomando responsabilidad personal y social* (TPRS), de Escartí, Pascual y Gutiérrez (2005), presenta exclusivamente una concordancia consensuada, no profundizando en los análisis de calidad del dato. También 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 de esta investigación obtienen mejores indicadores.

En cuanto a la herramienta presentada por Hernández Mendo, Díaz Martínez y Morales (2010), con el título, “*Herramienta observacional para evaluar las conductas prosociales en las clases de educación física*”, el abordaje realizado guarda características comunes en el análisis de calidad del dato. 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 intraobservadores e interobservadores, hasta los estudios de generalizabilidad de los datos (Blanco y Hernández-Mendo, 1998). Esta herramienta está construida para observar este mismo proceso en la ESO.

La herramienta “*ad hoc*” que presentamos permite obtener registros fiables del comportamiento del alumnado durante las sesiones de EF. La ventaja de este sistema es su fácil manejo, al mismo tiempo es exhaustivo y las categorías son mutuamente excluyentes. Esta herramienta ofrece al evaluador, investigador y/o a los profesionales de EF, un instrumento útil y sencillo para el desarrollo de investigaciones (evaluación de Programas, investigaciones cuasi-experimentales, observaciones, etc.) En consecuencia, se trata de un instrumento de registro adecuado al ámbito social de la EF en Primaria.

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, se hace preciso llevar a cabo una complementariedad metodológica (Anguera, 2004) y plantear programas y sus correspondientes evaluaciones que recojan esta propuesta (Planchuelo, 2008). Es viable pues 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). Se hace necesario pues, complementar otros estudios realizados en el ámbito del desarrollo moral con la utilización complementaria de diferentes metodologías (Cechini, Montero y Peña, 2003; Dias, Cruz y Danish; 2000; Vivó, 2001).

REFERENCIAS

- Anguera, M. T., Blanco, 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, 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].
- Anguera, M. T. (1979, abril). Observación de la conducta espacial. Comunicación presentada al VI Congreso de Nacional de Psicología, Pamplona, España.
- 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 de la Universidad de Murcia.
- Anguera, M.T. (1993). Proceso de Categorización. En M. T. Anguera (Ed.), *Metodología Observacional en la investigación psicológica. Vol I: Fundamentación*. Barcelona: PPU.
- Anguera, M. T. (2004). Posición de la metodología observacional en el debate entre las opciones metodológicas cualitativa y cuantitativa. ¿Enfrentamiento, complementariedad, integración? *Psicología en Revista*, 10(15), 13-27.
- Anguera, M. T. y Blanco, A. (2006). *¿Cómo se lleva a cabo un registro observacional?* Barcelona: ICE-UB.
- Anguera, M. T., Blanco, A. y Losada, 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. y Blanco, A. (2003). Registro y codificación del comportamiento deportivo. En A. Hernández-Mendo (Coord.), *Psicología del Deporte (Vol. II): Metodología* (pp. 6-34). Buenos Aires: Tulio Guterman.
- Ardá, A. (1998). *Análisis de los patrones de juego en fútbol a 7. Estudio de las acciones ofensivas*. Tesis Doctoral no publicada, Universidade da Coruña, La Coruña, España.
- Bakeman, R. y Quera, V. (1996). *Análisis de la interacción. Análisis secuencial con SDIS y GSEQ*. Madrid: Rama.
- Blanco, A. (1993). Fiabilidad, precisión, validez y generalizabilidad de los diseños observacionales. En M. T. Anguera (Ed.), *Metodología observacional en la investigación psicológica: Vol. 2: Fundamentación (2)* (pp.151-261). Barcelona: PPU.
- Blanco, A. (1997, septiembre). *Precisión en la evaluación de la Investigación Observacional*. Comunicación presentada al V Congreso de Metodología de las CC. Humanas y Sociales, Sevilla, España.
- Blanco, 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 (Eds.), *Medio Ambiente y Responsabilidad Humana* (pp. 579-583). A Coruña: Tórculo.
- Blanco, A., Castellano, 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, 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 no publicada, Universidad Autónoma de Barcelona, Barcelona, España.
- Bredemeier, B. (1991). Morality and sport for all. En P. Oja y R. Telama (Eds.), *Sport for all* (pp. 365-372). Amsterdam: Elsevier.
- Castellano, J. (2000). *Observación y análisis de la acción del juego en el fútbol*. Tesis Doctoral no publicada, Universidad del País Vasco, País Vasco, España.
- Castellano, J., Hernández-Mendo, A. y Haro, 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, J. A., Montero, J. y Peña, 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.

- Cronbach, L. J., Gleser, G. C., Nanda, H. y Rajaratnam, N. (1972). *The dependability of behavioural measurements: theory of generalizability for scores and profiles*. New York, NY: John Wiley.
- Cruz, J., Boixadós, M., Valiente, 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.
- 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, A., Pascual, C. y Gutiérrez, M. (2005). *Responsabilidad personal y social a través de la educación física y el deporte*. Barcelona: Editorial Grao.
- Escartí, A. y García, A. (1994). Factores de los iguales relacionados con la práctica y la motivación deportiva en la adolescencia. *Revista Psicología del Deporte*, 6, 35-51.
- Fleiss, J. L. (1981). *Statistical methods for rates and proportions*. New York, NY: Wiley.
- Garay, J. O., Hernández-Mendo, A. y 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.
- Garay, J. O., Hernández-Mendo, A. y Morales, 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.
- Gorospa, 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.
- Gutiérrez, M. y Escartí, 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*. Tesis Doctoral no publicada, Universidad de Santiago de Compostela, Santiago de Compostela, España.
- Hernández-Mendo, A., Díaz Martínez, F. y Morales, V. (2010). Construcción de una herramienta observacional para evaluar las conductas prosociales en las clases de educación física. *Revista de Psicología del Deporte*, 19(2), 305-318.
- Kohlberg, L. (1981). *Essays on Moral Development. vol. 1: The Philosophy of Moral Development*. San Francisco, CA: Harper & Row.

- Kohlberg, L. (1992). *Psicología del desarrollo moral*. Bilbao: Desclée de Brouwer.
- Lee, M. (1988). Values and Responsibilities in Children's Sports. *Physical Education Review*, 11, 19-27.
- Losada, J. L. y Arnau, J. (2000). Fiabilidad entre observadores con datos categóricos mediante el ANOVA. *Psicothema*, 12(Supl. 2), 335-339.
- Piaget, J. (1977). *El criterio moral en el niño*. Barcelona: Fontanella.
- Planchuelo, L. (2008). *Intervención y evaluación del desarrollo moral en educación física en primaria*. Tesis Doctoral no publicada, Universidad de Málaga, Málaga, España.
- Romance, T. J. (1984). *A Program to promote moral development through Elementary School Physical Education*. Tesis Doctoral no publicada, Oregon University, Eugene, Oregon.
- Vivó, P. (2001). *Desarrollo de valores y razonamiento moral a través de la educación física en educación secundaria*. Tesis Doctoral no publicada, Universidad de Valencia, Valencia, España.
- 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.

Manuscrito recibido: 31/05/2012

Manuscrito aceptado: 11/10/2012