

La mejora en la evaluación formativa de maestros de educación física través de un instrumento de metaevaluación didáctica

The improvement in the formative assessment on physical education teachers through an instrument for metaevaluation

Vicente Navarro Adelantado, Francisco Jiménez Jiménez

Universidad de La Laguna, España

Resumen

Se parte de la relevancia de saber manejar las claves de la evaluación para que ésta adquiera un valor formativo, así como de determinados rasgos de la evaluación formativa que constituyen esas claves. En este estudio se identifican cinco rasgos que conforman una vía de metaevaluación didáctica para la evaluación y mejora de los modelos de evaluación en educación física. Para comprobar su viabilidad, hemos llevado a cabo un estudio de casos con cuatro maestros de educación física, mediante un proceso de investigación-acción, situado entre una unidad didáctica inicial y otra experimental, donde en esta última se aplicaron las alternativas diseñadas en el proceso anterior. El estudio ha seguido un diseño test-pos test en el que se han utilizado el instrumento de metaevaluación didáctica y dos cuestionarios (maestros, y alumnado), con el fin de comprobar los cambios producidos, habiéndose constatado el valor de autodiagnóstico del instrumento y un cambio docente hacia postulados más formativos.

Palabras clave: evaluación formativa; metaevaluación; educación física; reflexión-acción.

Abstract

The present paper is based on the relevance of managing evaluation keys in order to reach a formative value and on how these keys lie on some given features of formative assessment. A self-diagnostic instrument is applied to the teachers that roots in five characteristics of formative evaluation in physical education, hence conforming a didactic metaevaluation way about their evaluation models. A case study about four physical education teachers is carried out through a research-action process between an initial and an experimental didactic units. In the latter, the alternatives designed in a precedent process are applied. The above mentioned metaevaluation instrument and two questionnaires (teaching staff and pupils) constitute the test-posttest design of the study. Such a design is aimed at checking the experimented changes and has verified the self-diagnostic value of the instrument and a teaching change to more formative postulates.

Key words: formative assessment; metaevaluation; physical education; action research.

Correspondencia/correspondence: Vicente Navarro Adelantado

Departamento de Didáctica de la Expresión Musical, Plástica y Corporal (Universidad de La Laguna)

E-mail: vnavarro@ull.es

Introducción

Mejorar es consustancial a la evaluación formativa, siendo una acción alcanzada por los maestros cuando éstos son capaces de encontrar en qué reside su propia mejora. De ahí que someter a valoración la propia evaluación sea un procedimiento forzoso para este fin, porque permite analizar los componentes que hacen formativa la evaluación que emplean los maestros.

Todo modelo de evaluación es susceptible de ser evaluado; así, la metaevaluación constituye una fase más de la estructura del diseño didáctico (Castillo y Cabrerizo, 2003, 133). De esta manera, al evaluar los modelos de evaluación en el ámbito educativo nos situamos en una metaevaluación didáctica (MD), porque “las ideas que se aplican a la evaluación son trasladables a la metaevaluación” (Santos y Moreno, 2004, 915). En nuestro caso, estamos preocupados por conocer en qué aspectos-clave ha de mejorar el autodiagnóstico de la evaluación en el proceso de enseñanza y aprendizaje. Por consiguiente, postulamos una función metaevaluatora dentro del sistema de evaluación formativa de educación física.

Ahora bien, el problema de la evaluación, en cuanto a su valor educativo y procedimiento de alcance transformador, pasa por la aplicación de la evaluación formativa con intención de mejora de los procesos educativos durante su desarrollo (Scriven, 1967; Taras, 2005). Ello supone reconocer la evaluación como una verdadera estrategia formadora (Bordas y Cabrera, 2001; Cabezas, González y Carpintero, 2009). De ahí que un modelo formativo de evaluación necesite mostrar coherencia y consistencia en distintos elementos didácticos capaces de diagnosticar el aprendizaje y la enseñanza. De modo que nos situamos dentro de la evaluación formativa *para* el aprendizaje (Hay, 2006; Birenbaum, Breuer, Cascallar, Dochy, Dori, Ridgway, y Wiesemes, 2006; S. Chappuis y J. Chappuis, 2008), interesados en cómo mejorar el sistema de evaluación de los maestros.

Concebimos la evaluación como un elemento sistémico a través del cual se puede diagnosticar el modelo didáctico de los maestros. En este sentido, los procesos colaborativos entre docentes facilitan afrontar con garantías el cambio docente (Elliot, 1990, 1993, 1994; McAlpine, Wewston, Beauchamp, Wiseman, y Beauchamp, 1999; McKernan, 1999; Schön, 1983, 1987; Smyth, 1991; Stenhouse, 1987), que supone poseer esta capacidad diagnóstica al reconocer indicadores y discutir sobre evidencias que constatan la presencia o ausencia de ellos. Una concepción más interactiva de las relaciones entre investigadores y maestros, nos hace entender que los procesos de indagación y reflexión sobre la práctica de la propia evaluación son un asunto que debe compaginar lo que se diagnostica como debilidad, en el desarrollo de la actuación docente, con el compromiso para transformar la enseñanza.

Nuestro trabajo alude a la metaevaluación en educación física desde el punto de vista didáctico, con el fin de que el maestro sea capaz de autodiagnosticar el alcance formativo de los modelos de evaluación que emplea en el aula y de saber cómo mejorarlos. El nivel didáctico de la metaevaluación supone situar el foco más allá de los estándares de calidad de la evaluación de programas (Castejón, 1995, 1997, 1998, 2007; De la Orden y Martínez de Toda, 1992; Perona y Castejón, 2001; Rodríguez-Campos, 2004; Scriven, 1967; Stake, 2006; Stufflebeam y Shinkfield, 1987) para señalar el trasfondo de la práctica docente y de los propios implicados en los procesos (Bilbao, 2011; Velázquez y Hernández, 2004, 16); también supone, intencionadamente, dejar a un lado al alumnado de este fenómeno de autodiagnóstico y de mejora, a pesar de reconocer su implicación sistémica (López, 2000; Santos y Fernández-Río, 2009).

La escasa presencia de la metaevaluación ha sido considerada como una de ‘las patologías de la evaluación de la educación física’ (Díaz, 2005). Se han realizado pocas investigaciones dirigidas al empleo de la MD en la fases interactiva y postactiva asociada a procesos de cambio docente de los maestros, realizada por ellos mismos, y a la vez atendiendo a su dimensión didáctica (F. Jiménez, Navarro y H. Jiménez, 2001; Jiménez y Navarro, 2008; López, 1999a). Estas últimas aportaciones centran su atención en la valoración que hace el profesorado sobre los modelos de evaluación del aprendizaje que emplea en la programación de aula.

Un concepto de evaluación formativa más amplio, y que alcanza al sentido del valor formativo y cómo hacerlo compartido, es lo que se viene haciendo desde la Red Nacional de Evaluación Formativa y Compartida, que profundiza —en la docencia universitaria— en las relaciones coherentes entre los procesos de enseñanza y aprendizaje, los criterios y los instrumentos de evaluación (López, 2009). En esta Red es el propio profesorado quién también realiza la valoración de su propio sistema de evaluación, y dentro de esta valoración incluye la opinión del alumnado acerca de las ventajas e inconvenientes que suscita la evaluación recibida, con el objeto de identificar las debilidades y potencialidades de su sistema de evaluación. En nuestro caso, la opinión del alumnado ha estado dirigida a contrastar los rasgos estructurales del sistema de evaluación aplicado.

Por tanto, la relación que nos interesa entre metaevaluación y evaluación formativa se concreta en la identificación de los rasgos que hacen posible que la evaluación cumpla una función formativa y que éstos sean identificados para el diagnóstico. Bajo este planteamiento, hemos intervenido con el grupo de maestros que han participado en un estudio generando un proceso reflexivo con el fin de debatir y asumir los aspectos de la evaluación que desembocan en el valor formativo de sus propuestas.

La discusión acerca de la evaluación formativa en educación física ha estado centrada, por una parte, en la identificación de un conjunto de rasgos integrados que hacen de ésta una actividad formadora (Blázquez, 1990; Chivite, 1999; López, 2009; Maccario, 1989; Ramos, del Villar y Moreno, 1998; Sebastiani, 1993; Siedentop, 1998; Velázquez y Hernández, 2004; Velázquez, 1991); y, por otro lado, abordando ‘criterios de calidad evaluativa’ (López, 1999a, 17), que sirvan como ‘criterios de metaevaluación educativa’ (López, 2000, 184), o como condiciones integradas de eficacia de la evaluación (Hay y Penny, 2009).

Desde estos planteamientos, hemos seguido un proceso de investigación-acción centrado en el empleo formativo de la evaluación. Este proceso, iniciado mediante una metaevaluación diagnóstica, se dirige a la elaboración de alternativas personales de evaluación (Rosales, 2000, 152) y todo ello, a través de un instrumento *ad hoc* que contiene cinco rasgos de la evaluación formativa y que permite reconocerlos en la práctica. Asumiendo que no es fácil cambiar los estereotipos establecidos por el profesorado de manera personal con relación a la función que éstos atribuyen a la evaluación (Nunziati, 1990, 51), hemos optado por un proceso de trabajo colaborativo que facilitara la labor del grupo estudiado. Los objetivos de esta investigación, son: conocer el alcance formativo que tienen los modelos de evaluación que emplean los maestros estudiados en el desarrollo de sus unidades didácticas, y valorar la incidencia de un instrumento de MD en el cambio de los modelos de evaluación empleados por los maestros.

Método

Este trabajo recoge el diagnóstico y transformación del modelo de evaluación de un grupo de cuatro maestros de educación física de Educación Primaria. El estudio se organizó a través de un diseño cuasiexperimental test-pos test, intercalando un proceso reflexivo colaborativo entre ambas pruebas

Esta investigación se abordó desde una perspectiva paradigmática de complementariedad (Cook y Reichardt, 1986; Flick, 2004), en el marco de un estudio de casos grupal en el que se integró un diseño test-pos test a través de una metodología selectiva (Arnau, Anguera y Gómez, 1990) y de un proceso de investigación-acción de carácter colaborativo (Sagor, 1992) (Figura 1).

El diseño fue descriptivo y cuasiexperimental comparativo, sin grupo de control, a través de un procedimiento de aplicación de una unidad de didáctica inicial (1ª fase), sin instrucción. Tras ésta se aplicó una segunda unidad didáctica experimental (3ª fase). Entre ambas pruebas se realizó un proceso colaborativo de investigación-acción que partió del problema de optimizar la función formativa de los modelos de evaluación de los maestros (2ª fase).



Fig. 1. Fases de la investigación

Con el fin de desvelar la incidencia del proceso de investigación-acción en la evolución de los modelos de evaluación y conducirlos hacia postulados más formativos, se aplicaron dos cuestionarios de opinión (test-pos test). Un cuestionario pretendía conocer la percepción del alumnado respecto a aspectos didácticos de la evaluación; el otro cuestionario se dirigió a que los maestros autodiagnosticaran el perfil didáctico de sus modelos de evaluación. Asimismo, se empleó un instrumento específico de MD (Jiménez y col., 2001), con el fin de contrastar el cuestionario de autodiagnóstico de los maestros. Para una mejor interpretación y comprensión del fenómeno estudiado, se realizó una triangulación (Denzin, 1978; Jick, 1979) con los datos obtenidos en ambas unidades de enseñanza de cada maestro mediante los cuestionarios y los datos provenientes del instrumento de MD.

Se filmaron todas las sesiones de educación física de cada una de las dos unidades didácticas desarrolladas por cada maestro, grabándose los discursos docentes en audio mediante un soporte inalámbrico de solapa (AZDEN), con el fin de disminuir los efectos perturbadores durante las clases. El contenido de las filmaciones se empleó para contrastar los resultados de los autodiagnósticos que se atribuían los maestros y para comprender el desarrollo didáctico de sus clases y sus cambios, según las fases de la investigación. También se llevó a cabo un análisis de las programaciones e instrumentos de evaluación empleados por los maestros, con el fin de comprobar el modelo didáctico seguido por cada uno de ellos.

Por consiguiente, este diseño metodológico trata de responder a la realidad de la enseñanza como hecho social y práctico, que reconoce la capacidad del maestro como agente transformador de su modelo didáctico. Por tanto, en coherencia con nuestra concepción paradigmática, el desarrollo de esta investigación complementa lo analítico y comparativo con lo procesual e interpretativo.

Participantes

La selección de los casos (E, O, JL, A) fue de carácter incidental. Se seleccionaron entre maestros especialistas de educación física, y dentro del ámbito territorial de la isla de Tenerife; todos los maestros fueron hombres, con más de diez años de experiencia profesional en educación física. De estos maestros, tres de ellos eran especialistas de educación física y uno, además, poseía la licenciatura de Pedagogía (O). La muestra del alumnado fue seleccionada por cada maestro, eligiendo uno de los cursos de tercer ciclo (11-12 años) que tenía a su cargo. El número de alumnos fue de 66, niños y niñas, que constituían los grupos normalizados de clase.

Instrumentos

a) Ficha de metaevaluación didáctica

Este instrumento (Tablas 1 y 2) se organiza según los rasgos que hemos considerado que son caracterizadores de la evaluación formativa.

Los rasgos actúan como dimensiones con las que se analiza cada uno de los instrumentos de evaluación que activa el maestro. Los criterios considerados son los siguientes:

- 1) Integración de la evaluación del aprendizaje y de la enseñanza en el sistema de evaluación.
- 2) Implicación activa del alumnado en los procesos de evaluación.
- 3) Uso de instrumentos de evaluación de forma continua durante los procesos de enseñanza y de aprendizaje.
- 4) Presencia de un flujo bidireccional constante de comunicación maestro-alumno respecto a la información obtenida en las actividades de evaluación.
- 5) Criterios de evaluación cognitivos, motrices y afectivos-sociales que convergen en el contenido de enseñanza.

Siguiendo estos rasgos, se diseñó la ficha de metaevaluación didáctica.

Tabla 1. Ficha de metaevaluación didáctica. El caso J. L. (unidad didáctica 1)

FICHA de META EVALUACIÓN DIDÁCTICA											
Maestro/a: J. L.			Centro:				U.D. 1 (Mini-Hockey)				
INSTRUMENTOS	EVALUACIÓN		IMPLICACIÓN		MOMENTOS de la EVALUACIÓN			FLUJO de COMUNICACIÓN	CRITERIOS de EVALUACIÓN		
	APR.	ENS.	A	P	Ini.	Pro.	Fin.		C	M	AF/S
Lista de control	√			√		√					√
Prueba motriz	√					√	√			√	

Nota: APR.= Aprendizaje; ENS.= Enseñanza; A= Alumnado; P= Profesorado; Ini.= Inicial; Pro.= Proceso; Fin.= Final; C= Ámbito cognitivo; M= Ámbito motor; A/F= Ámbito afectivo-social

Tabla 2. Ficha de metaevaluación didáctica. El caso J. L. (unidad didáctica 2)

FICHA de META EVALUACIÓN DIDÁCTICA											
Maestro/a: J. L.			Centro:				U.D. 2 (Mini-Tenis)				
INSTRUMENTOS	EVALUACIÓN		IMPLICACIÓN		MOMENTOS de la EVALUACIÓN			FLUJO de COMUNICACIÓN	CRITERIOS de EVALUACIÓN		
	APR.	ENS.	A	P	Ini.	Pro.	Fin.		C	M	AF/S
Registro anecdótico		√		√		√					√
Cuestionario de opinión	-		√		√	√		Masivo Individualizado			√
Lista de control	√		√			√	√	Masivo Individualizado		√	√
Prueba teórica de relación	√		√				√		√		

Nota: APR.= Aprendizaje; ENS.= Enseñanza; A= Alumnado; P= Profesorado; Ini.= Inicial; Pro.= Proceso; Fin.= Final; C= Ámbito cognitivo; M= Ámbito motor; A/F= Ámbito afectivo-social

En el 'cuestionario de opinión' se acredita la implicación del alumnado mediante 'autoevaluación'. La 'lista de control' se aplica al alumnado mediante un procedimiento de 'evaluación recíproca'.

Con este instrumento de MD se pudo conocer los rasgos de los modelos de evaluación utilizados por los maestros antes y después de la aplicación de las unidades didácticas correspondientes. La aplicación del instrumento pretendió, también, que los maestros tomarán conciencia de los límites y del alcance formativo de sus modelos de evaluación por medio del contraste de evidencias de la presencia o ausencia de los rasgos.

b) Cuestionario de autodiagnóstico didáctico para los maestros

Con este instrumento se sondea el perfil didáctico que los maestros atribuyen a su evaluación (Tabla 4). Este instrumento se elaboró específicamente para esta investigación. Está compuesto por 14 ítems con un formato de escala de cuatro alternativas (nada, poco, bastante y mucho, codificadas de 0 a 3), sus enunciados responden a cinco áreas o factores que se corresponden con rasgos caracterizadores de la evaluación formativa, coincidiendo alguna de estas dimensiones con las consideradas en la elaboración del cuestionario aplicado al

alumnado. Las dimensiones fueron: criterios de evaluación (ítems 5, 10, 11, 13); evaluación inicial (ítems 1, 2, 3, 7); evaluación del aprendizaje durante el proceso (8, 12); implicación del alumnado en la evaluación del aprendizaje (6, 9); evaluación de la enseñanza (4, 14).

La validación del instrumento ha sido realizada por dos expertos, ambos profesores universitarios e investigadores relacionados con la metodología de investigación, que constataron la coherencia interna entre los enunciados y los contenidos. Al tratarse de cuatro casos, se desestimó la utilidad de aplicar una prueba de consistencia interna del cuestionario.

c) *Cuestionario aplicado al alumnado acerca de aspectos didácticos de su aprendizaje y de su enseñanza.*

Con el empleo de este cuestionario se buscó conocer la opinión del alumnado acerca de cómo es evaluado, con el objeto de desvelar si los alumnos discriminan los cambios que se introducen en los modelos de evaluación (Tabla 5). Este cuestionario consta de 10 ítems, con un formato de escala de 4 alternativas (nada, poco, bastante y mucho, codificadas de 0 a 3); para evitar posibles sesgos en la respuesta del alumnado, el cuestionario fue aplicado por los investigadores. El anonimato de los alumnos se preservó con una clave (palabra o código secreto). El cuestionario está organizado en 4 dimensiones, las cuales son coincidentes con los rasgos caracterizadores de la evaluación formativa. Las dimensiones fueron: evaluación de la enseñanza (ítem 4); implicación del alumnado (ítems 3, 6, 9, 10); evaluación del momento inicial (ítems 1, 2); evaluación del proceso (ítems 5, 7, 8). El número total de ítems está justificado por la edad de los sujetos y la complejidad de los conceptos didácticos sobre los que se les cuestiona.

Este cuestionario ya había sido aplicado (Jiménez y col., 2001) a un grupo de 94 alumnos, con un índice de fiabilidad de 0.71, según el procedimiento de consistencia interna Alfa de Cronbach. Este índice se considera moderado; si bien, hay que tener en cuenta que este valor está altamente influido por el número total de ítems de la prueba, cuya consistencia podemos considerarla aceptable.

Proceso de investigación colaborativa

El grupo de investigación colaborativa, formado por los cuatro maestros y los dos investigadores, comienza a funcionar, como tal, al finalizar la primera fase de la investigación; tuvo como objetivo reconocer los fundamentos didácticos de la evaluación formativa y su aplicación, y promover los compromisos docentes necesarios para mejorar los modelos de evaluación. Se realizaron cinco sesiones, con una duración de entre dos y tres horas; se optó por un ciclo corto de sesiones debido a que se trataba de incidir sobre un aspecto didáctico concreto de la evaluación. El contenido de la sesiones incluyó comentarios de textos, debates acerca del valor pedagógico de la evaluación y sus correspondientes elementos didácticos, autodiagnóstico, visualización y comentarios de vídeos del mismo maestro estudiado impartiendo las clases, tomas de decisiones y elaboración de alternativas formativas de evaluación, éstas últimas con vista a la siguiente unidad didáctica.

Resultados

A continuación, se presentan los resultados en el siguiente orden: autodiagnóstico inicial y final de los maestros acerca de sus modelos de evaluación, realizado por medio de la ficha de MD (Tablas 1 y 2), y a través del cuestionario de opinión de los maestros (Tabla 4); seguidamente, los resultados del cuestionario de opinión del alumnado acerca de la evaluación que se le ha aplicado en las fases primera y tercera de la investigación (Tabla 5).

Autodiagnóstico inicial y final de los maestros

En la Tabla 3, se recoge de manera sistematizada el perfil del modelo de evaluación que empleó cada maestro en el pretest y en el pos test, respecto a los rasgos de la evaluación formativa (ficha de MD).

Tabla 3. Resultados comparativos de los modelos de evaluación empleados por los maestros, obtenidos a través de la aplicación del instrumento de metaevaluación

Indicadores de la metaevaluación didáctica	Casos pretest				Casos pos test			
	A	E	JL	O	A	E	JL	O
Integración de la evaluación del aprendizaje y de la enseñanza.				✓	✓	✓	✓	✓
Implicación activa del alumnado en los procesos de evaluación del aprendizaje.				✓	✓	✓	✓	✓
Implicación activa del alumnado en los procesos de evaluación de la enseñanza.					✓	✓		✓
Evaluación inicial.				✓	✓		✓	✓
Evaluación durante el proceso.	✓		✓	✓	✓	✓	✓	✓
Evaluación sumativa.			✓			✓	✓	✓
Flujo de comunicación.				✓		✓	✓	✓
Criterios de evaluación del aprendizaje: ámbito cognitivo.				✓	✓		✓	✓
Criterios de evaluación del aprendizaje: ámbito motor.	✓		✓	✓	✓	✓	✓	✓
Criterios de evaluación del aprendizaje: ámbitos afectivo-social.			✓	✓	✓	✓	✓	✓

a) *Integración de la evaluación del aprendizaje y de la enseñanza.*

Como se puede apreciar, en tres de los casos se produjo una transformación de sus modelos de evaluación, pasando a integrar en ellos la evaluación de la enseñanza junto con la evaluación del aprendizaje.

b) *Implicación activa del alumnado en los procesos de evaluación del aprendizaje y de la enseñanza.*

La falta de implicación inicial del alumnado en la evaluación de la enseñanza fue reconocida por los maestros y subsanada a través del proceso de reflexión colaborativa. Los maestros implicaron a su alumnado mediante cuestionarios de opinión acerca de la enseñanza. Respecto a la participación del alumnado en la evaluación de su aprendizaje, sólo el caso "O" lo consideró inicialmente mediante autoevaluación en el ámbito motor, y a través de evaluación recíproca para evaluar los ámbitos motor y afectivo-social. En el pos test, se aprecia que los maestros implicaron al alumnado en la evaluación de su aprendizaje mediante la evaluación recíproca, valorando mayoritariamente el ámbito motor y, en segundo término, el ámbito afectivo-social.

c) *Momentos de la evaluación.*

En cuanto a la evaluación inicial, nos encontramos con unos modelos de evaluación que no la consideran, en 3 de los 4 casos. En el pos test, estos datos se invierten. Sin embargo, esta evaluación inicial, en ningún caso, atiende a los tres ámbitos de desarrollo. La evaluación inicial se enfoca, prioritariamente, al ámbito cognitivo (prueba de múltiple opción y lista de control).

La evaluación realizada durante el proceso ya formaba parte del modelo de evaluación de los maestros, excepto en el caso de un maestro que no empleaba instrumentos para evaluar, y que sí los incorporó, junto a los demás, en el pos test. En cuanto a los ámbitos de desarrollo, encontramos una integración de lo cognitivo, motor y afectivo-social en tres de los cuatro casos en el pos test.

d) Flujo de comunicación.

En esta interacción docente-discente se sustenta el efecto formativo de la evaluación que hace posible la orientación del aprendizaje del alumnado. Los resultados muestran que la orientación del aprendizaje, que se genera al compartir con el alumnado la información recogida en los instrumentos de evaluación (ausente en el test), se integra en los modelos de evaluación de los maestros en tres de los cuatro casos.

e) Criterios de evaluación del aprendizaje.

Sólo uno de los casos tomó como referencia los tres ámbitos curriculares de formación (cognitivo, motor y afectivo-social). Por el contrario, tras el autodiagnóstico inicial se consigue que dos casos más integren los diversos ámbitos en su modelo de evaluación. El ámbito motriz ha sido atendido de manera preferente, como se observa en los datos de la Tabla 3, y en los resultados del cuestionario de autodiagnóstico de los maestros (ítem 13, $\bar{x}=1,50$; $\bar{x}=2,00$).

En general, los resultados apuntan a que el empleo del instrumento de MD permitió a los maestros identificar las carencias de sus modelos de evaluación. La información diagnóstica aportada por este instrumento activó en los maestros la búsqueda de alternativas con el fin de dotar a sus modelos de evaluación de un mayor valor formativo. Esta transformación se produjo en todos los maestros, incluso en aquél que partía de un modelo de evaluación más completo.

Finalmente, la información obtenida del empleo del instrumento aplicado para el autodiagnóstico inicial y final de los maestros (Tabla 2) pone de manifiesto el cambio que se ha producido en los modelos de evaluación utilizados.

Resultados del cuestionario aplicado a los maestros.

Los resultados se sometieron a la prueba de los signos de Wilcoxon de diferencias de medianas para muestras relacionadas. Estos resultados obtenidos aluden, por tanto, a la diferencias entre los ítems test-pos test de las unidades didácticas 1 y 2. Como puede observarse en la Tabla 4, se aprecian mejoras en el pos test en la mayoría de los ítems (2, 4, 6, 7, 8, 9, 11, 12, 13), lo que puede interpretarse como que los maestros perciben su cambio en la evaluación formativa; sin embargo, sólo un ítem, referido al manejo de los instrumentos de evaluación del aprendizaje por parte del alumnado (ítem 6), muestra un valor relevante ($p=.059$). En general, los maestros han sobrestimado su desempeño didáctico con elevadas valoraciones de partida en los ítems del test, y esta cuestión, a nuestro juicio, ha impedido la aparición de más diferencias estadísticas en el resto de los ítems.

Los resultados del ítem 6 confirman un efecto positivo producido cuando tres de los cuatro maestros pasaron de no implicar a los alumnos en la evaluación de su aprendizaje a implicarlos activamente (Tabla 3). Con ello, se puede interpretar que este indicador ha sido el que ha generado una conciencia más clara entre los maestros respecto a la transformación sufrida en sus modelos de evaluación.

Tabla 4. Resultados del análisis del cuestionario de autodiagnóstico didáctico para los maestros

ÍTEM	Momento	N	\bar{x}	desv. típica	Z	p
1. ¿Das a conocer a tus alumnos los objetivos que pretendes conseguir antes de un contenido nuevo?	pretest	4	1.50	0.58	.000	1
	pos test	4	1.50	0.58		
2. ¿Solicitas la opinión de tus alumnos para ayudarte en la mejora de tu planteamiento didáctico?	pretest	4	1.00	0.00	-1.000	.317
	pos test	4	1.25	0.50		
3. ¿Crees que la manera que tienes de evaluar a tus alumnos permite que ellos establezcan niveles personales de mejora?	pretest	4	1.75	0.50	-1.000	.317
	pos test	4	1.50	0.58		
4. ¿Consideras la opinión de tus alumnos para evaluar cómo impartes la clase?	pretest	4	1.75	1.26	-.577	.564
	pos test	4	2.00	0.82		
5. ¿Evalúas contenidos afectivo-sociales?	pretest	4	2.50	0.58	-1.732	.083
	pos test	4	1.75	0.96		
6. ¿Empleas instrumentos para que el alumnado pueda evaluar su aprendizaje?	pretest	4	0.50	0.58	-1.890	.059
	pos test	4	1.75	0.50		
7. ¿Realizas alguna iniciativa específica para conocer el nivel inicial del grupo antes de comenzar un contenido nuevo?	pretest	4	1.25	0.50	-1.414	.157
	pos test	4	1.75	0.50		
8. ¿Crees que las actividades de clase permiten al alumnado comprobar si mejoran en su aprendizaje?	pretest	4	1.75	0.50	-1.000	.317
	pos test	4	2.00	0.82		
9. ¿Promueves que unos alumnos participen en la evaluación de otros?	pretest	4	1.00	0.82	-.816	.414
	pos test	4	1.50	0.58		
10. ¿Evalúas contenidos cognitivos?	pretest	4	1.00	0.82	.000	1
	pos test	4	1.00	0.00		
11. ¿Enuncias en cada unidad didáctica criterios que orienten la evaluación del aprendizaje?	pretest	4	1.25	0.50	-1.000	.317
	pos test	4	1.50	0.58		
12. ¿Comentas con tus alumnos la información que obtienes mediante instrumentos de evaluación para orientar su aprendizaje?	pretest	4	1.75	0.50	-1.000	.317
	pos test	4	2.00	0.00		
13. ¿Evalúas contenidos motrices?	pretest	4	1.50	0.58	-1.414	.157
	pos test	4	2.00	0.00		
14. ¿Crees que tu autoevaluación serviría para mejorar tu enseñanza?	pretest	4	2.25	0.50	.000	1
	pos test	4	2.25	0.50		

*** $p=.000$

** $p<.005$

Resultados del cuestionario aplicado al alumnado.

Igualmente, respecto al cuestionario anterior, se aplica la prueba de los signos de Wilcoxon de diferencias de mediana para muestras relacionadas, ya que se trata de un diseño de una sola muestra pretest-pos test.

Como se observa en la Tabla 5, un conjunto amplio de ítems muestran significaciones estadísticas (ítem 1, $p=.003$; ítem 2, $p=.003$; ítem 3, $p=.000$; ítem 6, $p=.003$; ítem 9, $p=.000$; ítem 10, $p=.003$), y con una tendencia importante en otro de ellos (ítem 4, $p=.006$). Los ítems con significación del primer grupo, excepto los ítems 1 y 2 relacionados con la dimensión *evaluación inicial*, forman parte de una misma dimensión: *implicación del alumnado*. Esta concentración en este rasgo de la evaluación formativa evidencia la sensibilidad del alumnado a los cambios producidos en las propuestas docentes realizadas en este ámbito. Por otra parte, el hecho de que los ítems 5, 7 y 8, pertenecientes a la dimensión *evaluación del proceso*, no obtengan significación estadística ha podido estar condicionado por los altos valores de puntuación que presentaban estos ítems en el pretest.

Tabla 5. Resultados del análisis del cuestionario del alumnado

ÍTEM	Momento	N	\bar{x}	desv. típica	Z	p
1. Antes de comenzar una nueva actividad o tema ¿te dice tu maestro qué es lo nuevo que vas a aprender?	pretest	66	2.44	0.84	-2.992	.003**
	pos test	66	2.82	0.52		
2. Antes de empezar una nueva actividad o tema ¿tu maestro hace preguntas a la clase para ver qué es lo que ya saben?	pretest	66	1.50	1.11	-2.996	.003**
	pos test	66	1.89	0.91		
3. Antes de que tu maestro te evalúe ¿sabes hasta dónde has aprendido o mejorado?	pretest	66	1.61	0.99	-3.879	.000***
	pos test	66	2.27	0.73		
4. ¿Les pregunta el maestro acerca de cómo da su clase?	pretest	65	0.73	0.97	-2.729	.006
	pos test	65	1.15	1.06		
5. ¿Tu maestro te da indicaciones o pautas para que mejores en tu aprendizaje?	pretest	65	2.66	0.66	-.620	.535
	pos test	65	2.73	0.57		
6. ¿El maestro permite que te evalúes tú mismo?	pretest	65	0.55	0.83	-2.964	.003**
	pos test	65	1.02	0.96		
7. ¿Crees que te ayuda a aprender la manera de evaluar en las clases que tiene tu maestro?	pretest	65	2.06	0.92	-1.198	.231
	pos test	65	2.22	0.87		
8. Con las actividades que realizas en clase ¿tú mismo te puedes dar cuenta de si mejoras o no?	pretest	66	2.45	0.74	-1.339	.181
	pos test	66	2.61	0.67		
9. Mejora tu aprendizaje si tus compañeros o compañeras participan en tu evaluación?	pretest	66	0.94	1.02	-5.034	.000***
	pos test	66	2.03	1.00		
10. Crees que mejoraría el aprendizaje de tus compañeros o compañeras si tú participarías en su evaluación?	pretest	66	1.14	1.17	-2.990	.003**
	pos test	66	1.82	1.10		

*** $p=.000$

** $p<.005$

Discusión

La triangulación de los resultados obtenidos, mediante el empleo del instrumento de autodiagnóstico y los cuestionarios de opinión, aplicados a los maestros y a su alumnado, evidencian que los rasgos de la evaluación formativa que hemos manejado aportan claves para la MD. Nuestra investigación ha desvelado que la dimensión *implicación del alumnado en la evaluación de su aprendizaje* es el rasgo estudiado de la evaluación formativa que trae más consecuencias en la percepción de los alumnos respecto al cambio docente, y en el autodiagnóstico de los maestros acerca de sus modelos de evaluación. Es decir, que la implicación del alumnado en la evaluación del aprendizaje debería formar parte como indicador más revelador de cualquier modelo de metaevaluación educativa, tal y como refrendan las significaciones encontradas en los ítems del cuestionario del alumnado: ítem 3, $p=.000$; ítem 6, $p=.003$; ítem 9, $p=.000$; ítem 10, $p=.003$

En este sentido, estamos de acuerdo con Nunziati (1990), López (2000), Earl (2003), Hay, (2006), Birenbaum y col. (2006), S. Chappuis y J. Chappuis (2008) cuando sostienen una evaluación a partir del alumno y tomando como referencia su aprendizaje. En nuestro caso, esta implicación se ha detectado en la conciencia del alumnado en conocer hasta dónde ha aprendido y mejorado (ítem 3), en poder evaluarse él mismo (ítem 6), y en compartir la evaluación con sus compañeros (ítem 6, 9 y 10).

Además, en nuestra investigación se confirma la relación entre la evaluación formativa y algunas competencias básicas curriculares (Currículo de Educación Primaria. BOE 173, 2007. p. 31488), como son las competencia de *aprender a aprender* con el ítem 3 del cuestionario del alumnado, la competencia de *autonomía e iniciativa personal* con los ítems 6, 9 y 10 del mismo cuestionario. En esta línea, la evaluación formativa resulta una verdadera estrategia formadora para el alumnado (Bordás y Cabrera, 2001).

Tradicionalmente, la educación física centró la participación didáctica del alumnado en la presentación y evaluación o autoevaluación de la tarea mediante estilos de enseñanza, como la enseñanza recíproca y la autoevaluación de tareas (Mosston y Ashworth, 1996). Sin embargo, hemos constatado que es necesario no confundir la tarea como referencia de la participación en la evaluación con las actividades de evaluación vinculadas al seguimiento del desarrollo de capacidades en el proceso de enseñanza y aprendizaje, y, de esta manera, no se reduce el alcance didáctico de la participación del alumnado en la evaluación de su aprendizaje. La evaluación de una unidad didáctica se ha de centrar en el seguimiento de determinadas capacidades, que se manifestarán en diversos momentos y en múltiples tareas. En nuestro caso, la participación didáctica que hemos considerado se circunscribe al alto valor de la implicación del alumnado en la evaluación de su aprendizaje a lo largo de la unidad didáctica. Este aspecto nos ha permitido discriminar con precisión la incidencia formativa de los modelos de evaluación empleados por los maestros.

Respecto a los momentos de la evaluación, el cambio docente ocurrido al incorporar la evaluación inicial a los modelos de evaluación ha sido reconocido favorablemente de forma significativa por el alumnado (ítem 1, $p=.003$; ítem 2, $p=.003$). Este cambio repercute en la posibilidad que ofrece al maestro de adecuar el proceso de enseñanza a las necesidades reales de aprendizaje del alumnado. En este caso, el proceso colaborativo de reflexión apoyado en la MD ha provocado una reformulación de los momentos de la evaluación a favor de la evaluación inicial y la evaluación durante el proceso.

La triangulación ha desvelado también dos divergencias surgidas entre los participantes en el acto educativo. Estas divergencias son, por un lado, la percepción que los alumnos tienen de la intervención docente del maestro y la percepción que el propio maestro tiene acerca de su actuación; y, por otra parte, la segunda divergencia surge de la falta de correspondencia entre lo manifestado por los maestros en el cuestionario y la ficha de MD, acerca de los modelos de evaluación empleados en las unidades didácticas 1 y 2.

El fenómeno de la divergencia en la percepción entre profesorado y alumnado ya es advertido en educación física por Pieron (1999, 242) como algo frecuente en los procesos de enseñanza y aprendizaje, lo que, en nuestra opinión, señala la importancia de considerar ambas percepciones en los diseños de investigación educativa. En nuestro caso, los maestros han sobrevalorado su modelo de evaluación habitual, cuando hacen esta valoración a través de un cuestionario; ello se comprueba en la inexistencia de diferencias estadísticamente significativas en los ítems del cuestionario aplicado a los maestros. Mientras, esta diferencia test-pos test sí es percibida por el alumnado, reflejándose en las diferencias significativas de la mayoría de los ítems del cuestionario del alumnado (Tabla 5). Esta valoración realizada por los alumnos confirma la existencia de un cambio en los modelos de evaluación empleados por sus maestros, antes y después del proceso de reflexión colaborativa.

Esta falsa percepción de la propia enseñanza mostrada por los maestros, se puede interpretar por un exceso de confianza de éstos amparados en su larga experiencia profesional, en su capacidad de retención de la información de los procesos, y en la relativización de sus planteamientos didácticos. El proceso de reflexión colaborativa ha conseguido concienciar a

los maestros de la relevancia de compartir con el alumnado la información recogida con los instrumentos de evaluación, con el objeto de orientar el aprendizaje del alumnado. Consideramos que el empleo de instrumentos es una técnica sólida para asegurar el flujo de comunicación maestro-alumno(s) con una finalidad orientadora y de mejora. Porque, si bien los maestros con experiencia son capaces de retener alguna información no reflejada en el instrumento con el cual activar la orientación del aprendizaje, esta información fragmentada, de no ser registrada, limita la función orientadora de la evaluación.

Ahora bien, la bondad o la falsa creencia de los maestros acerca de su percepción de la incidencia formativa de su modelo de evaluación, desaparece cuando éstos realizan su autodiagnóstico mediante la ficha de MD, tal y como se desprende en nuestro estudio. Este instrumento de autodiagnóstico aporta a los maestros las claves operativas con las que reconocer los límites formativos del modelo de evaluación empleado y activar el cambio en su intervención docente. El cambio no obedece al número de instrumentos introducidos en las clases, sino a que éstos cumplan con las funciones de la evaluación formativa. Nuestra afirmación se constata en todos los maestros estudiados, en cómo ellos han abordado la funcionalidad de los instrumentos de evaluación aplicados, y no en el aumento de instrumentos. De hecho, los maestros no son del todo conscientes de cuáles son las intenciones de sus propios modelos de evaluación, de cómo los están aplicando y de sus debilidades formativas, hasta que no los someten a un autodiagnóstico de metaevaluación por medio de algún instrumento, como ya comprobamos con anterioridad (Jiménez y col., 2001).

El instrumento de MD ha desvelado verdaderos criterios de calidad respecto a la evaluación formativa, en cuanto a utilidad, valor ético y precisión (López, 2000). La utilidad se ha concretado en el efecto transformador que el autodiagnóstico ha provocado en los modelos iniciales y deficitarios de evaluación. El valor ético del procedimiento de uso del instrumento se ha mostrado en el proceso de reflexión que ha provocado en los maestros y que ha traído consigo la mejora de su intervención docente, haciéndolos más responsables de los beneficios formativos que aporten sus modelos de evaluación en el alumnado, y respetando cada proceso personal. La precisión de este instrumento se ha acreditado mediante la identificación y la operatividad de los cinco rasgos caracterizadores de la evaluación formativa en educación física, lo que ha permitido a los maestros identificar con claridad las debilidades y fortalezas de sus modelos de evaluación. Y lo que es más importante, orientando su cambio docente.

Por último, estimamos que nuestro trabajo, desde el punto de vista de la ‘transferibilidad’ (Santos y Moreno, 2004, 926), y aun resultando irreplicable en cuanto al proceso cualitativo seguido con los maestros, contiene en el proceso colaborativo aspectos epistemológicos y prácticos transferibles, puesto que los maestros se han cuestionado qué grado de coherencia interna y problemas posee la evaluación formativa en una pedagogía de la educación física, y cómo aplicar los planteamiento que han ido asimilando a futuros contextos curriculares o de cambio de alumnado y centro escolar.

Conclusiones

El proceso seguido confirmó, con las debidas precauciones dado el número de maestros estudiados, el beneficio para el conocimiento práctico del *valor de lo colectivo* (Martínez, 2008, 8), y generó una serie de alternativas que fueron aplicadas y contrastadas a través de la propia valoración de los maestros y de la opinión del alumnado. Todos los maestros estudiados, independientemente de su nivel de partida, evidenciaron un cambio favorable y desarrollaron una actividad más completa desde el punto de vista formativo, a través de sus modelos de evaluación, con relación a los procesos de aprendizaje y enseñanza. En esta

transformación ha jugado un papel relevante el empleo de un instrumento específico de metaevaluación didáctica.

Con relación al objetivo acerca del alcance formativo de los modelos de evaluación que emplean los maestros estudiados, hemos constatado las deficiencias iniciales y la falta de identificación coherente del sistema de evaluación del que se partía.

Con relación al segundo objetivo acerca de la incidencia del instrumento de MD en el cambio de los modelos de evaluación empleados por los maestros, hemos comprobado que:

- El autodiagnóstico, a través de los cinco rasgos considerados en el instrumento de metaevaluación didáctica, y mediado por un proceso de reflexión colaborativo, ha aportado a los maestros las claves operativas con las que mejorar la calidad formativa de sus modelos de evaluación.
- La dimensión *implicación del alumnado en la evaluación de su aprendizaje* resulta el indicador metaevaluador más fiable respecto a otros en el contexto didáctico, actuando como un espejo entre las opiniones acerca de este indicador de los maestros y el alumnado.

Referencias bibliográficas

- Arnau, J.; Anguera, M. T., y Gómez, J. (1990). *Metodología de la investigación en ciencias del comportamiento*. Murcia: Universidad de Murcia.
- Bilbao, A. (2011). Teoría y práctica del valor estratégico de la evaluación del aprendizaje en la formación básica de maestros: una experiencia de metaevaluación a cargo de estudiantes. *Actas de las Jornadas Interuniversitarias de Innovación Docente*, 16 y 17 de junio (pp. 167-172). Barcelona: Universidad de Deusto, Universidad Pontificia de Comillas, y Universitat Ramon Llull.
<http://www.url.edu/jornadasinnovaciondocente/docs/librojornada.pdf>.
(Consultada: 18/07/2011).
- Birenbaum, M.; Breuer, K.; Cascallar, E.; Dochy, F.; Dori, Y.; Ridgway, J., & Wiesemes, R. (2006). A learning Integrated Assessment System. In: EARLI Series of Position Papers. R. Wiesemes, G. Nickmans A. (Eds.), *Educational Research Review*, 1(1), 61-67.
http://edu.technion.ac.il/chemicaleducation/judy/publications/no12_Assessment%20for%20learning%202006.pdf (Consultada: 12/09/2011)
- Blázquez, D. (1990). *Evaluar en Educación Física*. Barcelona: INDE.
- Bordas, M. I., y Cabrera, F. A. (2001). Estrategias de evaluación de los aprendizajes centrados en el proceso, *Revista Española de Pedagogía*, 218 enero-abril. 25-48.
- Cabezas, D.; González, C., y Carpintero, E. (2009). Evaluar la evaluación: cuestionario sobre prácticas de evaluación, *eduPsykné*, 8 (1), 51-61.
- Castejón, F. J. (1995). Construcción y utilización de indicadores en la Educación Física Escolar, *Revista Española de Educación Física y Deportes* 2(4), 20-26.
- Castejón, F. J. (1997). La evaluación de la programación en Educación Física: el panel de análisis de contenido en Educación Física (PACEF56), *Revista Española de Educación Física y Deportes* 4(4), 13-21.
- Castejón, F. J. (1998). Variables del PACEF56 para la evaluación de la programación en Educación Física. *Revista ASKESIS*, colaboraciones libres, 2.
- Castejón, F. J. (2007). *Evaluación de programas en ciencias de la Actividad Física*. Madrid: Síntesis.

- Castillo, S., y Cabrerizo, J. (2000). *La evaluación educativa hoy. Formación y práctica*. Vol. I. Formación. Madrid: UNED.
- Chappuis, S., & Chappuis, J. (2008). The Best Value in Formative Assessment. *Educational Leadership*, 65 (4), 14-18.
- Chivite, M. (1999). *Evaluación escolar en educación física: una interpretación en torno al proceso reflexivo del maestro en sus decisiones de aula*. Tesis doctoral. Universidad de Zaragoza.
- Cook, T. D., y Reichardt, C. S. (1986). *Métodos cuantitativos y cualitativos en investigación evaluativa*. Madrid: Morata.
- De la Orden, A., y Martínez de Toda, M. J. (1992). Metaevaluación educativa, *Bordón*, 43 (4), 517-527.
- Denzin, N. K. (1978). *The Research Act. A Theoretical Introduction to a Sociological Methods*. New York: McGraw Hill.
- Díaz, J. (2005). *La evaluación formativa como instrumento de aprendizaje en Educación Física*. Barcelona: INDE.
- Earl, L. M. (2003). *Assessment as learning: using classroom Assessment to maximize student learning*. Thousand Oaks: Corving Press.
- Elliot, J. (1990). *La investigación acción en educación*. Madrid: Morata.
- (1993). *El cambio educativo desde la investigación-acción*. Madrid: Morata.
- (1994). *La investigación-acción en educación*. Madrid: Morata.
- Flick, U. (2004). *Introducción a la investigación cualitativa*. Madrid: Morata.
- Jick, T. D. (1979). Mixing Qualitative and Cuantitative Methods: Triangulation in Action. *Administrative Science Quaterly*, 24 (4), 602-611.
- Jiménez, F.; Navarro, V., y Jiménez, H. (2001). Incidencia de la metaevaluación en la evaluación formativa de la educación física, *Revista Interuniversitaria de Psicología de la Educación. Evaluación e intervención psicoeducativa*, 6 y 7 (I), 259-273.
- Jiménez, F., y Navarro, V. (2008). Evaluación formativa y metaevaluación en educación física: dos estudios de casos colectivos en las etapas de educación primaria y secundaria. *Revista Española de Educación Física y Deportes*, 9, 13-25.
- Hay, P. (2006). Assessment for learning. In D. Kirk, D. Macdonald y M. O'Sullivan (Eds.), *The Hand-book of Physical Education*, (pp. 312-325). London: Sage.
- Hay, P., & Penny, D. (2009). Proposing conditions for assessment efficacy in physical education. *European Physical Education Review*, 15 (3), 389-405.
- López, V. (1999a) (coord.). *Educación física evaluación y reforma*. Segovia: Diagonal.
- López, V. (1999b). *Prácticas de evaluación en Educación Física: estudios de casos en Primaria, Secundaria y Formación del Profesorado*. Valladolid: Universidad de Valladolid.
- López, V. (2000). *Evaluación compartida*. Sevilla: Publicaciones MCEP.
- López, V. (2009) (coord.). *Evaluación formativa y compartida en Educación Superior. Propuestas técnicas, instrumentos y experiencias*. Madrid: Narcea.
- Maccario, B. (1989). *Teoría y práctica de la evaluación de las actividades físicas y deportivas*. Buenos Aires: Lidium.
- Martínez, J. (2008). El olvido de la investigación-acción en el asesoramiento docente y la innovación educativa. *Profesorado. Revista de currículum y formación del profesorado*, 12, (1). <http://www.ugr.es/~recfpro/rev121.ART3.pdf> (Consultado: 20/03/2010)

- McAlpine, L.; Wewston, C.; Beauchamp, C.; Wiseman, C., & Beauchamp, J. (1999). Building a metacognitive model of reflection. *Higher Education*, 37, 105-131.
- McKernan, J. (1999). *Investigación-acción y currículum*. Madrid: Morata.
- Mosston, M., y Ashworth. S. (1993). *La enseñanza de la Educación Física. La reforma de los estilos de enseñanza*. Barcelona: Hispano Europea.
- Nunziati, G. (1990). Pour construire un dispositif d'évaluation formatrice. *Cahiers Pédagogiques*, 280, 47-64.
- Perona, M., y Castejón F. J. (2001). La utilización del PACEF62 en la evaluación de la programación de la Educación Física en la ESO. *Revista Española de Educación Física y Deportes*, 4, 10-21.
- Pieron, M. (1999). *Para una enseñanza eficaz de las actividades físico-deportivas*. Barcelona: INDE.
- Ramos, L. A.; del Villar, F., y Moreno, P. (1998). La evaluación en el proceso de formación del profesorado de Educación Física. *Revista de Educación Física*, 70, 5-10.
- Rodríguez-Campos, L. (2004). Metaevaluation: Theoretical and Practical Guidelines for Evaluating Evaluations. *International Journal of Learning*, 11, 1527-1543.
- Rosales, C. (2000). *Evaluar es reflexionar sobre la enseñanza*. Madrid: Narcea.
- Sagor, R. (1992). *How to conduct collaborative Action Research*. Alexandria: ASCD.
- Santos, M. Á., y Moreno, T. (2004). ¿El momento de la metaevaluación educativa? *Revista Mexicana de Investigación Educativa*, 23 (9), 913-931. http://www.oei.es/evaluacioneducativa/momento_metaevaluacion_educativa_santos_guer ra.pdf (Consultada: 20/04/2011)”
- Santos, L., y Fernández-Río, J. (2009). El 'Cuaderno de Bitácora' de Educación Física. Elemento central dentro de una propuesta de metaevaluación. *Retos. Nuevas Tendencias en Educación Física, Deporte y Recreación*, 16. 92-96. http://www.retos.org/numero_15/retos%2016-18.pdf (Consultado: 15/5/2010).
- Scriven, M. (1967). The methodology of evaluation. En Tyler, R.; Gagne, R. y Scriven, M. (Eds.), *Perspectives of currículo evaluation* (pp. 39-83). Chicago: Rand McNally.
- Sebastiani, O. (1993). La evaluación de la Educación Física en la Reforma educativa. *Apunts Educació Física y Sports*, 31, 17-26.
- Siedentop, D. (1998). *Aprender a enseñar la educación física*. Barcelona: INDE.
- Smyth, J. (1991). Una pedagogía crítica de la práctica en el aula. *Revista de Educación*, 294, 275-300.
- Schön, D.A. (1983). *The Reflective Practitioner: How professional think in action*. London: Temple Smith [Trad. cast. *profesional reflexivo: Cómo piensan los profesionales cuando actúan*. Barcelona: Paidós, 1998.]
- (1987) *Educating the reflective practitioner*. San Francisco: Jossey-Bass. [Trad. cast. *La formación de profesionales reflexivos. Hacia un nuevo diseño de la enseñanza y el aprendizaje en las profesiones*. Barcelona: Paidós/MEC, 1992.]
- Stake, R. E. (2006). *Evaluación comprensiva y evaluación basada en estándares*. Barcelona: Graó.
- Stenhouse, L. (1987). *La investigación como base de la enseñanza*. Madrid: Morata.
- Stufflebeam, D. L. y Shinkfield, A. J. (1987). *Evaluación sistemática*. Madrid: Paidós MEC.
- Taras, M. (2005). Assessment –Summative and Formative– some theoretical reflections. *British Journal of Education Studies*, 53 (4), 466-478.

Navarro, V.; Jiménez, F. (2012). La mejora en la evaluación formativa de maestros de educación física través de un instrumento de metaevaluación didáctica. *RICYDE. Rev. int. cienc. deporte.* 27(8), 63-79. doi: 10.5232/ricyde2012.02705

Velázquez, R. (1991). La evaluación en Educación Física. *Cuadernos de Pedagogía*, 198, 26-29.

Velázquez, R., y Hernández, J. L. (2004). Evaluación en educación y evaluación del aprendizaje en educación física. En Velázquez, R. y Hernández, J. L. (coords.), *La evaluación en Educación Física* (pp. 11-48). Barcelona: Graó.