

José Devís Devís,
Profesor del IVEF,
Departamento de Didáctica, Expresión
Musical, Plástica y Corporal.
Carmen Peiró Velert,
Agregada de Bachillerato,
Generalitat Valenciana.

EVALUACIÓN DE PROGRAMAS: UN PROGRAMA DE EDUCACIÓN FÍSICA Y SALUD

Resumen

La proliferación de programas sociales sobre actividad física y salud hace de su evaluación un importante elemento del progreso social. En este artículo se comenta la evaluación cualitativa de un pequeño programa de educación física y salud desarrollado en varias clases de primero de BUP, después de un previo apoyo teórico en el que se presenta el carácter problemático de la evaluación, y se hace una breve revisión de las aproximaciones y modelos más importantes.

Palabras clave: evaluación de programas, ejercicio físico, salud, evaluación cualitativa, condición física.

Introducción: la evaluación de programas sociales

En nuestra vida cotidiana, las personas hacemos constantemente juicios de valor sobre acontecimientos, comportamientos y cosas. En el mundo laboral y profesional, los juicios se hacen para conocer, de una forma más sistemática y profunda, el valor y calidad de nuestro trabajo. La evaluación es, en líneas generales, un proceso sistemático de investigación en el que se ofrecen juicios de valor sobre algún tipo de fenómeno a partir de algún tipo de criterio.

La evaluación, según House (1980), descende de la modernización. Esta se libera de la tradición y provoca el cambio de una realidad social incuestiona-

ble (tradicional) por otra en la que todo puede cuestionarse y cambiarse. En el mundo moderno las oportunidades de elección aumentan considerablemente, pero el hacer elecciones también se convierte en un problema. De ahí que se proponga la evaluación como una ayuda para la elección. Si además los programas de servicios sociales proliferan en cualquier área de la actividad humana como la educación, la rehabilitación, el deporte, la recreación o la salud pública, la evaluación de programas sociales se convierte en un importante elemento del progreso social.

Tal y como apunta Rohs (1988) la evaluación de programas "está típicamente asociada al grado en que un programa es efectivo, adecuado, bueno o malo, valioso o inestimable, y a lo apropiadas o inapropiadas que pueden resultar ciertas acciones, procesos o resultados con respecto a cómo lo percibe la persona que utiliza la información de la evaluación" (p. 113).

Estos programas pueden definirse en sentido estricto o amplio, pueden ser grandes o pequeños proyectos y continuos o alternos. En cualquier caso, un programa social requiere la presencia de muchos recursos (humanos, conceptuales, técnicos y económicos) y afecta a muchas personas, lo que indica el elevado costo en tiempo y dinero, y en ocasiones un costo político (Angulo, 1988). Asimismo, en el desarrollo de un programa confluyen multitud de intereses provenientes de la institución patrocinadora del programa, de los profesionales que lo desarrollan, de los clientes, del evaluador/a y de otros grupos sociales que acaban por convertir la evaluación en un elemento problemático.

El carácter problemático de la evaluación de programas

El carácter problemático de la evaluación de programas se evidencia claramente con el conjunto de aspectos generales con que Kemmis y Stake (1988) caracterizan a la investigación evaluativa, y que pasamos a comentar brevemente.

- Cualquier perspectiva evaluativa se dirige u orienta a cuestiones relacionadas con algún criterio de calidad. Pero el problema está en que la calidad depende de las percepciones particulares de cada persona y de algún tipo de tradición social y cultural.
- La información que genera la investigación evaluativa debe orientarse a la interpretación y la comprensión de aquello que se evalúa. No se trata tanto de producir datos como de ordenarlos y presentarlos de forma que los lectores/as puedan interpretarlos y comprenderlos, es decir, ayudar a situar aquello que se evalúa dentro de la estructura de significados y valores de quien lo lee para obtener una idea de su calidad. Pero durante ese proceso, el lector y la lectora también evalúan la propia evaluación, esto es, las interpretaciones y juicios del evaluador/a.
- Los juicios de la evaluación se orientan, de alguna manera, a la acción. Adquieren valor porque ayudan a tomar decisiones, generalmente para mejorar lo que se evalúa. Pero esto no significa que se tomen decisiones de forma automática, sino que el proceso de evaluación debe invitar a que la gente adopte una actitud crítica. En palabras de los autores, invitan a "pensar sobre lo que es y

- debe ser, a pensar sobre las relaciones entre lo real y lo posible" (p. 9).
- d) La evaluación está inevitablemente relacionada con cuestiones de poder y en última instancia es siempre una cuestión "política". No existe una vinculación inocente o neutra entre la información y la toma de decisión. Lo que se evalúa se localiza "en un contexto de comprensión en el que se entrecruzan biografías personales, formas culturales, posiciones ideológicas y, haciendo de todo ello un objeto de comprensión, nos invita a pensar cómo pueden ser las cosas de otra manera" (p. 9).
- e) La materia o tema objeto de la evaluación es un aspecto importante del proceso, ya sea un programa recreativo o comunitario, un libro, un programa de ejercicio físico y salud, un currículum o la Reforma del Sistema Educativo. La materia permite comprender lo que se evalúa dentro de su contexto, aunque también puede ser visto en relación con otros contextos similares o diferentes.

Estos aspectos nos vienen a recordar que en la evaluación de un programa se aplica una jeraquía de valores donde los grupos de intereses rivalizan entre sí. Por lo tanto, las decisiones prácticas y metodológicas del evaluador/a además de ser "una opción ética involucra al evaluador en el proceso político" (Gimeno y Pérez, 1985, p.423). De ahí que algunos autores vean la evaluación como una forma de persuasión que es más creíble que cierta y más variable que prescriptiva. Por todo ello, parece conveniente hacer una breve incursión en las diferentes formas de persuasión que podemos encontrarnos en una sociedad pluralista.

Principales aproximaciones y modelos a la evaluación de programas

Existen formas muy variadas de evaluación, pero de entre ellas pueden identificarse ciertas aproximaciones, pers-

pectivas o tendencias fundamentales. House (1980) habla de una aproximación "utilitaria" y otra "pluralista". La primera se sustenta en una epistemología objetivista que se alcanza con la utilización de instrumentos "científicamente objetivos" como tests y cuestionarios. Intenta llegar a un único juicio que sea de utilidad para todos los programas en general y su finalidad última es la predicción. La segunda se sustenta en una epistemología subjetivista que considera que la verdad está relacionada con la formación y la experiencia de las personas. La evaluación está intencionalmente vinculada al contexto y los resultados de la evaluación se interpretan dentro del mismo. Su finalidad última es comprender a los participantes del programa.

Angulo (1988) se refiere a dos perspectivas evaluativas que se corresponden epistemológicamente con las aproximaciones de House y que denomina "perspectiva experimental y transaccional". La experimental pretende demostrar la eficacia de un programa, puesto que éste es el modo que se considera "racionalmente" válido para la toma de decisiones en política social, mientras que la perspectiva transaccional pretende ser "una plataforma para la argumentación, la crítica y la participación ciudadana en su progreso y desarrollo" (p. 194).

Por otra parte, MacDonald (1985) presenta tres tipos de investigación evaluativa a partir de una clasificación política, es decir, de acuerdo con los intereses a los que sirve: la evaluación burocrática, la autocrática y la democrática. La burocrática constituye un servicio incondicional a las agencias gubernamentales que mayor control poseen sobre los recursos. El evaluador/a acepta los valores de las autoridades o patrocinadores y ofrece la información que les ayude a realizar los objetivos de su política. Sus técnicas de estudio deben ser creíbles a los políticos y no deben ofrecerse a la crítica pública. Los conceptos clave de este tipo de evaluación son "servicio", "uti-



lidad" y "eficacia", y su concepto justificador "la realidad del poder". La evaluación autocrática es un servicio condicional, puesto que ofrece una validación externa a la política gubernamental a cambio de que acepten sus recomendaciones. El evaluador/a se concentra en los temas que presentan mayor valor y actúa como consejero experto. Sus técnicas producen pruebas científicas, porque su poder es la comunidad investigadora. Los conceptos clave de su evaluación son "los principios" y "la objetividad", y su concepto justificador la "responsabilidad" profesional. Por último, la evaluación democrática es un servicio de información a la comunidad sobre las características del programa. La evaluadora o el evaluador reconoce el pluralismo de valores, busca la representación de intereses diferentes en su formulación del tema e intercambia información entre grupos que desean conocimientos recíprocos. También garantiza el secreto a los informantes. Los conceptos clave son "confidencia", "negociación" y "accesibilidad", y su concepto justificador "el derecho de saber". Recientemente, Weiss (1989) se refiere a cinco tendencias dentro del campo de la evaluación. La primera es la ten-

dencia tradicional, positivista y cuantitativa, que se corresponde con la epistemología objetivista de House y que se caracteriza por la verificación científica y la predicción. El resto de tendencias se corresponderían con una epistemología subjetivista, pero con diferencias entre ellas. Así tenemos que la segunda es la tendencia cualitativa del crítico de arte (literario, de cine o de pintura), del jurista o del columnista periodístico. La tercera se caracteriza por la reacción a la concepción de experto que se atribuye al evaluador en las tendencias anteriores. El resurgimiento de la investigación/acción le ofrece la legitimidad suficiente para considerar que los prácticos, los profesionales ligados a una actividad práctica, poseen la experiencia y conocimiento suficiente como para evaluar sus propios programas. La cuarta tendencia es aquella que ve a la investigación evaluativa como un proceso educativo en la que el evaluador anima al cliente a abrirse a nuevos conocimientos sobre la base de la negociación cotidiana. Utiliza metodología cualitativa de investigación y recoge aspectos de la tendencia anterior así como de la evaluación democrática de MacDonald. No obstante, Kemmis considera que, aunque en la evaluación democrática el experto puede colaborar con los participantes, estos últimos juegan un papel más representativo que realmente participativo en el proceso de la evaluación. Por último, Weiss señala una tendencia proveniente de la filosofía en la Escuela de Frankfurt y conocida como teoría crítica. La evaluación se entiende como una actividad emancipadora que permitirá comprender cómo determinados aspectos de la vida de las personas están mediatizados por un sistema social en el que los recursos y el poder se distribuyen desigualmente.

Cada una de las aproximaciones que hemos presentado posee uno o varios tipos básicos de evaluación que hemos denominado "modelos" y pueden verse representados en la tabla 1. Cada



modelo posee sus propios defensores, su fundamentación, su propia historia de evaluaciones realizadas según su diseño básico, y sirven a distintos propósitos, audiencias e intereses. Con alguna excepción, utilizan técnicas de investigación social provenientes, fundamentalmente, del mundo de los negocios y la educación. Cualquier modelo puede ser apropiado o inapropiado dependiendo de las circunstancias y su aplicación y la correspondiente validez de los supuestos en los que se fundamenta. En realidad, son idealizaciones o tipos ideales de evaluación que se moldearán en función de las contingencias y condiciones de su uso, pero son importantes herramientas para situarnos dentro de este campo.

La evaluación de un programa de educación física y salud

Los programas sociales de ejercicio físico y salud no cuentan con una gran tradición en el contexto español. En los años ochenta, con la popularización del jogging y la gimnasia de mantenimiento surgieron programas privados y públicos de este estilo, pero con grandes

dosis de intuición y sin prestar mucha atención a su evaluación. A lo sumo, se siguieron planteamientos objetivistas ligados a la realización de tests de condición física, importados del campo de la educación física, que poseían multitud de lagunas metodológicas y acabaron convirtiéndose en simple rutina, e incluso perdiendo su pretendido papel evaluador. Además, la mayoría de los tests utilizados seguían las pautas de las grandes baterías de los años sesenta, propuestas por la AAHPER o cualquier otra prestigiosa asociación, y que según Pate (1989) se dirigen al rendimiento físico y atlético y no al ejercicio físico relacionado con la salud.

Aunque en la última década se han buscado relaciones entre la condición física/ejercicio físico y la salud, especialmente en los países de habla inglesa, la forma dominante de evaluación de este tipo de programas se corresponde con una perspectiva experimental, que en palabras de Godin y Shephard (1983) es la única "suficientemente rigurosa para proporcionar una evaluación categórica de la efectividad de los programas de promoción" (p. 105). Dentro de esta perspectiva el tipo más característico de evaluación es la "referida a normas". Es decir, aquella que establece exteriormente, por medio de una escala elaborada estadísticamente, unas referencias o niveles que generalmente corresponden al percentil 50 y que permiten comparar los resultados de un individuo con los resultados de la población o grupo al que pertenece. Es una forma de evaluación próxima al modelo de análisis de sistemas que más fielmente representa la concepción de la evaluación como pura medida.

Sin embargo, los tests relacionados con la salud deberían seguir, según Simons-Morton et al. (1988), la evaluación "referida a criterios" (próxima a un modelo de objetivos) porque pueden apreciarse los cambios y la mejora de un mismo alumno/a en cada prueba, y marca el nivel mínimo por el que un determinado atributo de la condición física se le relaciona con la salud. No

Modelo	Propósito	Proponente	Metodología	Resultado	Preguntas típicas	Problemas
Análisis sistemas	Relaciones causa y efecto; cuantificación de variables	Rivlin, Burnham	Planificación económica y gestión, coste y beneficio, tratamiento estadístico	Eficiencia y feedback	¿Se han alcanzado los efectos deseados? ¿Qué programa es el más eficiente?	Énfasis cuantitativo; obvia valores; no humanista
Objetivos conducta	Objetivos pre-establecidos; cuantificación variables	Tyler, Popham, Provus	Objetivos de conducta y test en referencia a criterios	Productividad y responsabilidad	¿El programa consigue los objetivos?	Desacuerdo en criterios y medidas del logro
Toma de decisión	Metas generales y criterios	Alkin, Stufflebeam	Cuestionarios y entrevistas	Efectividad control de calidad y elección	¿Es efectivo el programa o alguna parte?	Desprecia los intereses de las minorías
Sin referencia a objetivos	Reducir los efectos tendenciosos	Scriven	Análisis lógico y control de tendencias	Elección del cliente y utilidad social	¿Cuáles son todos los efectos?	Incomodidad psicológica de los participantes
Crítica artística	Valoración por experto conocedor de la materia	Eisner	Revisión crítica	Valoración subjetiva	¿Aumenta la apreciación crítica de la audiencia?	Diversidad visiones y poco desarrollo metodológico
Revisión profesional	Juicios de profesionales con experiencia	North Central Association, Henry	Revisión crítica por informes personales, entrevistas y cuestionarios	Aceptación profesional	¿Cómo valoran los profesionales el programa?	Desacuerdo en los criterios y la selección de los profesionales expertos
Tipo legal	Pros y contras del programa	--	Procedimientos de tipo legal	Resolución	¿Cuáles son los argumentos a favor y en contra?	Selección del jurado y de los procedimientos
Negociación	Retrato de lo acontecido	Stake, Parlett y Hamilton, y MacDonald y Walker	Estudio de casos, entrevistas y observación participante	Comprensión de la diversidad de visiones	¿Qué les parece el programa a los participantes y al evaluador?	Conflictos entre participantes y difícil análisis cualitativo
Investigación en la acción	Autoevaluación	Lewin, Stenhouse, Elliot, y Kemmis y McTaggart	Estudio de casos, entrevistas y observación activa	Valorar para promover el cambio	¿La autoreflexión y crítica han provocado cambios?	Gran compromiso de todos los participantes
Teoría crítica	Atención a desigualdades sociales	Apple, Lundgren, Kallos y Paulston	Crítica social desde posición teórica determinada	Crítica a los distintos modelos de evaluación	¿Quién se beneficia del programa?	No utiliza procedimientos claros

Tabla 1. Principales modelos para la evaluación de programas

Fuente: Elaborado a partir de House (1980), Howe (1980), Pérez (1985) y Kemmis y Stake (1988)

obstante, reconocen la existencia de dificultades para desarrollar criterios científicos debido a la falta de estudios longitudinales. Aún así, proponen que los criterios para los tests de condición/ejercicio físico y salud deberían basarse en: (a) el conocimiento de unos niveles de condición física acordes con unas capacidades funcionales aceptables para la infancia y adolescencia; y (b) el conocimiento de unos niveles de condición física consistentes con la evitación de los problemas de salud a largo plazo que se asocian a la inactividad física.

Las recientes baterías de tests, ligadas a los programas norteamericanos de condición física relacionada con la salud (por ejemplo, el *Fitnessgram* y el *Physical Best*), siguen una evaluación referida a criterios cuyos niveles mínimos, tal y como señalan Corbin y Pangrazi (1992), los establecen los expertos a partir de la evidencia científica disponible. Pero no existe información definitiva sobre ninguna de las pruebas y atributos de la condición física relacionada con la salud que indique cuál es el nivel mínimo necesario para la salud (Looney y Plowman, 1990; Corbin, 1987).

Tanto la evaluación referida a normas como la referida a criterios en los pro-

gramas de ejercicio físico y salud posee gran cantidad de problemas. Seefeldt y Vogel (1989) recogen un grupo importante de ellos:

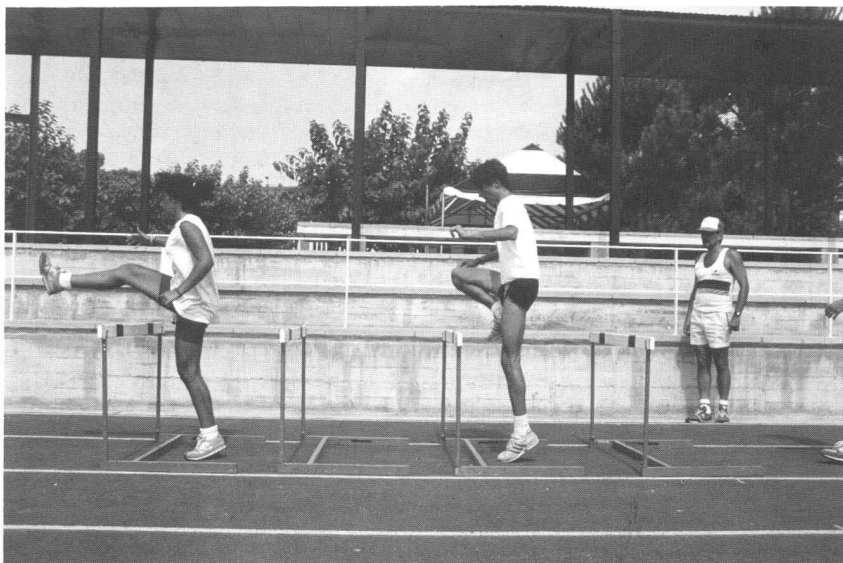
1. No hay acuerdo al definir la condición física y su relación con la salud, de ahí la dificultad en seleccionar las pruebas y baterías adecuadas.
2. Se pretende valorar la salud relacionada con la condición física sin ofrecer indicadores directos del estado de salud de los participantes (por ejemplo, presión sanguínea, respuesta cardíaca o lípidos en sangre).
3. Se enfatiza un comportamiento saludable y a la vez se anima a que los clientes realicen unos tests que pueden implicar violaciones de una conducta saludable (por ejemplo, máxima realización de abdominales en un minuto, pruebas de estiramientos balísticos o de resistencia aeróbica sin una adecuada preparación).
4. Se pide participación en actividades físicas dentro de los programas de salud y a la hora de su valoración se realiza cuantitativamente.
5. No existe evidencia científica para relacionar los niveles de los distintos componentes físicos de la salud en poblaciones jóvenes con su estado de salud cuando sean adultos.

6. No se proporciona la evidencia necesaria para decir que alcanzando los niveles que indican los tests se obtengan aquellos beneficios saludables con los que pretenden estar asociados.

7. No se demuestra que las mejoras alcanzadas en la batería de tests tengan mejoras en el estado de salud de los participantes.

Además, los resultados en estos tests se ven enormemente influidos por la carga genética de la persona que los realiza, los posibles cambios en las condiciones de realización, la variación motivacional de los sujetos al realizarlos y la habilidad específica que puede desarrollarse haciéndolos (Fox y Biddle, 1986). Los más críticos manifiestan que estas formas de evaluación implican un modelo individualista, competitivo y mecanicista de salud que no tiene en cuenta las desigualdades sociales, políticas y económicas de las personas que los realizan (Ineson y Sim, 1989).

Tal vez debería pensarse en evaluar el proceso (la actividad física) en lugar del producto asociado al resultado (la condición física), o ambos a la vez. Debemos tener presente que los verdaderos beneficios saludables de la actividad física se alcanzan haciendo actividad física y no buscando resultados, y que los niveles de actividad para obtener beneficios saludables no son necesariamente los mismos que para mejorar u obtener niveles aceptables de condición física (Meredith, 1988; Biddle y Biddle, 1989; Looney y Plowman, 1990). Desde nuestro punto de vista, una evaluación realizada únicamente mediante tests físicos es, cuanto menos, una evaluación reduccionista de un programa social de ejercicio físico y salud. Además del resultado o producto físico existen muchos factores personales y sociales, y otros relacionados con el contexto en que se desarrolla el programa, que también deben tenerse en cuenta.



Algunas cuestiones relacionadas con la evaluación de nuestro programa

A continuación comentaremos la evaluación de un programa de ejercicio físico y salud que se desarrolló dentro de las clases de educación física de primero de BUP (1988-1989), por lo que lo definimos como un programa de "educación física y salud".

Debemos tener en cuenta que la educación física escolar está considerándose en los últimos años como un área de actuación preferente para la promoción de la salud en niños/as y jóvenes por su influencia en el conocimiento, las actitudes y los valores que ayudan a mantener el interés y la participación en actividades físicas (Simons-Morton et al., 1987; Bar-Or, 1987; Simons-Morton et al., 1988; y Sallis y McKenzie, 1991). Como hemos visto, las evaluaciones realizadas siguen preferentemente una perspectiva experimental (ver también la propuesta de Mateo, 1992 en este monográfico), aunque algunos proyectos aislados han seguido una perspectiva pluralista/transaccional/cualitativa como el estudio de casos de Kirk (1986) y últimamente el proyecto colaborativo de Hutchinson et al. (1990).

En nuestro caso, y tal y como veremos a continuación, se trata de una evaluación cualitativa enclavada en un modelo de negociación. No es una evaluación de encargo, puesto que responde al interés conjunto de la profesora y el evaluador por conocer cómo se desarrolló el programa a partir de las opiniones de sus participantes. La colaboración entre el investigador y la persona que tomaba las decisiones después de la evaluación estaban claramente dirigidas por el propósito de mejorar el valor educativo del programa (Devís y Peiró, 1992).

Para comentar esta evaluación nos servirán de guía general las consideraciones que plantea Rohs (1988) respecto a la evaluación cualitativa.

El programa y sus propósitos

El programa era una modificación del Exercise Challenge desarrollado por McGeorge (1988) dentro del HEA-

Health and Physical Education Project de Gran Bretaña. Se caracterizaba por:

- Una duración de seis semanas, después de las explicaciones pertinentes.
- Los alumnos/as debían escribir y desarrollar su propio programa a partir de ciertas consideraciones expuestas por la profesora y la ayuda de material curricular, y teniendo en cuenta sus posibilidades, limitaciones y gustos.
- El material curricular consistía en unos esquemas para poder seguir sus planificaciones y en el artículo "El infarto y el ejercicio físico" (Michener, 1984) que debían comentar por escrito.
- Se firmaba un contrato individual con la profesora comprometiéndose a desarrollar el programa y dedicar un tercer día fuera del horario de la asignatura.
- La profesora no jugaba un papel central y no controlaba las decisiones de los alumnos/as. Se dedicaba a sugerir, dar guías básicas de orientación y revisar al inicio y mitad del programa los planes de los alumnos/as.
- Se pasaron unas pruebas tests al principio, a mitad y final del programa para ver la mejora del alumno/a. Se creía que era un importante elemento motivador y que serviría para reajustar los planes a mitad del programa.

Los propósitos del programa no eran objetivos conductuales que debieran comprobarse con su puesta en práctica sino, más bien, afirmaciones que sirvieran de guía y orientación durante todo el proceso de evaluación. Entre ellos destacaban:

- La participación activa de los alumnos y alumnas en la realización de actividades físicas, buscando promocionar estilos de vida activos y no buscar rendimiento ni grandes proezas físicas.

- La participación de las alumnas-os en el proceso de implementación del programa. Es decir, se trata de desarrollar un programa propio, o lo que es igual, tomar decisiones en el proceso de enseñanza.
- Favorecer la conciencia crítica de los alumnos/as en el contexto del ejercicio físico y la salud.

Elección del método y las técnicas de recogida de datos

Optamos por una metodología cualitativa o naturalística (Lincoln Guba, 1985) porque considerábamos que era el mejor modo de describir y comprender en profundidad un hecho particular, como era la implementación del programa de salud dentro de las clases de educación física. Además, la veíamos más acorde con los propósitos del programa y la evaluación.

Cuando se elige un enclave particular y el enfoque se centra en las perspectivas y experiencias de los participantes, el investigador cualitativo puede analizar lo que los participantes piensan y sienten como realidades significativas que facilitan la comprensión de su conducta en el contexto.

Las técnicas de recogida de datos utilizadas fueron las entrevistas informales, las observaciones y el análisis de documentos, provenientes tanto de los alumnos/as como de la profesora. Según Olson (1980), estas técnicas cualitativas están más al alcance de las habilidades y necesidades de los prácticos/profesores y permiten recoger datos para realizar una evaluación formativa y sumativa (sobre estos conceptos ver Blázquez, 1992; y Sebastiani, 1992 en esta revista).

Análisis e interpretación de los datos

El análisis e interpretación de los datos los realizó el evaluador, teniendo en cuenta que un investigador cualitativo, más que recoger datos para probar hipótesis preconcebidas, los analiza inductivamente y busca categorías y temas que surgen de esos datos. El análisis se comentó con la profesora.

Este análisis reveló cuatro áreas temáticas que nos ayudaron a presentar la información: 1. problemas o limitaciones contextuales; 2. implicación y participación del alumno/a; 3. diferencias de sexo y mejora; y 4. prevención y estilo de vida. Para no extendernos en el comentario de cada una de ellas, los resultados los presentaremos en la tabla 2. en base a pros y contras hallados. Estos resultados son importantes para mejorar los futuros programas de educación física y salud.

Utilización de la información y los juicios para el futuro

La evaluación del programa arrojó importantes consideraciones para la profesora. A partir del informe la profesora tomó importantes decisiones para futuros programas, entre los que se encuentran:

- Supresión total de los tests porque no se adecúan a un planteamiento que busca la participación y la experiencia en lugar del rendimiento. Probablemente la poca relación entre el trabajo realizado en clase y el resultado de los tests se deba a que éstos no son los instrumentos adecuados para evaluar a las alumnas y alumnos que participan en este programa.
- Introducción de actividades deportivas como una actividad más del programa, pero sin ser la única o la que más domine en una sesión de actividad.
- Introducción de actividades grupales para la participación conjunta de chicos y chicas.
- Atender de forma especial la conciencia crítica de los participantes al programa, introduciendo nuevas lecturas y alguna discusión en clase sobre temas problemáticos de ejercicio físico y salud.
- La realización de próximos programas de este estilo, pero dentro de un proyecto más ambicioso que ocupe unos dos trimestres del curso. El pro-

PROS	CONTRAS
<ul style="list-style-type: none"> – Negociación positiva sobre la realización de una tercera sesión fuera del horario de la asignatura, aunque con algunas críticas. – Gran importancia y valoración de los alumnos-as sobre su participación en la toma de decisiones: percibido como un signo de autonomía y responsabilidad personal. – Mejor opinión de los alumnos-as de estas clases de educación física en comparación con sus experiencias anteriores. – Adaptación del programa a las particularidades de los alumnos-as. – Negociación positiva, especialmente con los chicos, para incluir actividades deportivas entre las posibles a elegir (con limitaciones). – Enorme participación de las chicas en actividades físicas, especialmente de carácter más individual. – Apreciación del ejercicio físico como un elemento de prevención y estilo de vida sano, sin llegar a percibirse como la panacea que cura todos los males. 	<ul style="list-style-type: none"> – Problemas ligados a las malas condiciones de las instalaciones y al horario escolar que limitaban la cantidad de actividades físicas a realizar y el número de sesiones semanales (quedó en tres), e imposibilitaban la utilización de espacios municipales alternativos. – Críticas de los alumnos-as a las pruebas-tests porque no se relacionan con el trabajo realizado en el transcurso de las clases. También la profesora ve poca relación. Fue especialmente evidente en la prueba de resistencia, puesto que algunos alumnos-as eran capaces de mejorar la carrera continua media hora o más sin observar grandes cambios en la comparación de los tests. – La lectura y el trabajo sobre el artículo propuesto, a pesar de haberse entendido, no incide claramente sobre la conciencia crítica de los alumnos-as.

Tabla2. Pros y contras en la implementación del programa

grama de seis semanas sería la culminación de un proceso que incluyera en sus inicios aspectos básicos que ayuden a realizar y comprender mejor las relaciones entre ejercicio físico y salud (conocimiento práctico sobre seguridad en la realización de ejercicios, toma de pulsaciones y márgenes saludables de actividad física, etc.).

Conclusión

En este artículo, que versa sobre la evaluación de programas, se ha presentado el carácter problemático de la

evaluación después de hacer una breve introducción al tema. Asimismo, se ha realizado una pequeña revisión de las aproximaciones y modelos a la evaluación de programas para situarnos en un campo apenas explorado en el contexto español de la educación física.

Posteriormente, se analiza el problema de la evaluación en relación a los programas de ejercicio físico y salud, donde predomina una perspectiva experimental, objetivista y centrada en la medida de tests de condición física. Esta perspectiva, además de poseer multitud de problemas para evaluar realmente la salud vinculada a la práctica

de la actividad física, responde a una simple acepción biologicista o mecanicista puesto que olvida el carácter holístico de la salud, en la que intervienen aspectos psicológicos y sociales o personales y contextuales, y la concibe como un compartimento estanco que puede medirse con tests físicos. Presentar, finalmente, una pequeña evaluación cualitativa comentada no es algo casual, ya que responde a una postura determinada que entiende a la evaluación "como un instrumento real y fecundo, no artificial ni administrativo, de perfeccionamiento interno" (Gimeno y Pérez, 1985, p.424) de los programas sociales.

Bibliografía

- ANGULO, J.F. "Evaluación de programas sociales: de la eficacia a la democracia" *Revista de Educación*, 1988, núm. 286, pp. 193-206.
- BAR-OR O.A. "Commentary to children and fitness: a public health perspective" *Research Quarterly for Exercise and Sport*, 1987, v. 58, núm. 4, pp. 304-307.
- BIDDLE, S.; BIDDLE, G. "Health-related fitness for the primary school" En A. Williams (ed.) *Issues in Physical Education for the Primary Years*, East Sussex: The Falmer Press, 1989, pp. 54-75.
- CORBIN, C.B. "Youth fitness, exercise and health: There is much to be done" *Research Quarterly for Exercise and Sport*, 1987, v. 58, núm. 4, pp. 308-14.
- CORBIN, C.B.; PANGRAZI, R.P. "Are american children and youth fit?" *Research Quarterly for Exercise and Sport*, 1992, v. 63, núm. 2, pp. 96-106.
- DEVIS, J.; PEIRO, C. Exercise and health in a Spanish P.E. curriculum: a modified programme of "The Exercise Challenge". En T. Williams; L. Almond; A. Sparkes (eds.) *Sport and Physical Activity. Moving towards excellence*, London: E. and F.N. Spon, 1992, pp. 418-428.
- FOX, K.R.; BIDDLE, S.J.H. "Health related fitness testing in schools: introduction and problems of interpretation" *Bulletin of Physical Education*, 1986, núm. 22, pp. 54-64.
- GIMENO, J.; PÉREZ, A. "La evaluación didáctica." En J. Gimeno; A. Pérez (eds.) *La enseñanza: su teoría y su práctica*, Madrid: Akal, 1985, pp. 421-425.
- GODIN, G.; SHEPHARD, R.J. "Physical fitness promotion programmes: effectiveness in modifying exercise behavior" *Canadian Journal of Applied Sport Sciences*, 1983, v. 8, núm. 2, pp. 104-113.
- HOUSE, E.R. *Evaluating with validity*. London: Sage Publications, 1980.
- HOWE, C.Z. "Models for evaluating public recreation programs: What the literature shows" *Journal of Physical Education and Recreation*, 1980, October, pp. 36-38.
- HUTCHINSON, G.E.; FREEDSON, P.S.; WARD, A.; RIPPE, J. "Ideal to real. Implementing a youth fitness program" *Journal of Physical Education, Recreation and Dance*, 1990, agosto, pp. 52-58.
- INESON, A.; SIM, J. "Testing times: fitness testing and health" *Radical Community Medicine*, 1989, otoño, pp. 5-10.
- KEMMIS, S.; STAKE, R. *Evaluating curriculum*. Victoria: Deakin University Press, 1988.
- KIRK, D. "Health related fitness as an innovation in the physical education curriculum." En J. Evans (ed.) *Physical education, sport and schooling*, London: The Falmer Press, 1986, pp. 167-181.
- LINCOLN, Y.S.; GUBA, E.G. *Naturalistic inquiry*. London: Sage Publications, 1985.
- LOONEY, M.A.; PLOWMAN, S.A. "Passing rates of american children and youth on the Fitnessgram criterion-referenced physical fitness standards" *Research Quarterly for Exercise and Sport*, 1990, v. 61, núm. 3, pp. 215-223.
- MACDONALD, B. "La evaluación y el control de la educación." En J. Gimeno; A.I. Pérez (eds.) *La enseñanza: su teoría y su práctica*, Madrid: Akal, 1985, pp. 467-478.
- MATEO, J. "A la salut mitjançant l'exercici físic. Cal mesurar?" *Apunts Educació Física i Esports*, 1993, monográfico de evaluación.
- MCGEORGE, S. *The Exercise Challenge*. Loughborough: LUT, 1988.
- MEREDITH, M.D. "Activity or fitness: Is the process or the product more important for Public Health?" *Quest*, 1988, núm. 40, pp. 180-86.
- MICHENER, J. "El infarto y el ejercicio físico" *Diario 16*, 1984, suplemento "Ciencia".
- OLSON, E.G. "Program portrayal: a qualitative approach to recreation program evaluation." *Journal of Physical Education and Recreation*, 1980, octubre, pp. 41-42.
- PATE, R.R. "The case for large-scale physical fitness testing in american youth." *Pediatric Exercise Science*, 1989, núm. 1, pp. 290-94.
- PEREZ, A.I. "Modelos contemporáneos de evaluación." En J. Gimeno; A.I. Pérez (eds.) *La enseñanza: su teoría y su práctica*, Madrid: Akal, 1985, pp. 426-449.
- ROHS, F.R. "The qualitative researcher as evaluator." En J.P. Goetz; J. Allen (eds.) *Qualitative research in education. Substance, methods, experience*, Athens: The University of Georgia, 1988, pp. 113-122.
- SALLIS, J.F.; MCKENZIE, T.L. "Physical Education's Role in Public Health." *Research Quarterly for Exercise and Sport*, 1991, v. 62, núm. 2, págs 124-37.
- SEBASTIANI, E.M. "L'avaluació de l'educació física en la reforma educativa: un segon nivell de concreció" *Apunts Educació Física i Esports*, 1993, monográfico de evaluación.
- SEEFELDT, V.; VOGEL, P. "Physical fitness testing of children: a 30-years history of misguided efforts?" *Pediatric Exercise Science*, 1989, núm. 1, pp. 295-302.
- SIMONS-MORTON, B.G.; O'HARA, N.M.; SIMONS-MORTON, D.G.; PARCEL, G.S. "Children and fitness: a public health perspective" *Research Quarterly for Exercise and Sport*, 1987, v. 58, núm. 4, pp. 295-302.
- SIMONS-MORTON, B.G.; PARCEL, G.S.; O'HARA, N.M.; BLAIR, S.N.; PATE, R.R. "Health related physical fitness in childhood" *Annual Review of Public Health*, 1988, núm. 9, pp. 403-425.
- WEISS, J. "Evaluation as subversive educational activity" En G. Milburn, I.F. Goodson; R.J. Clark (eds.) *Re-interpreting curriculum research: images and arguments*, London: The Falmer Press, 1989, pp. 121-131.