

## Herramienta de evaluación de los programas deportivos para la promoción de la salud

Leonor Gallardo\*, Javier Sánchez-Sánchez\*\*, Ferrán Calabuig\*\*\*,  
Esther Ubago-Guisado\*, Pablo Burillo\*\*\*\*, Álvaro Fernández-Luna\*\*\*\*,  
Jorge García-Unanue\*\*\*\* y José Luis Felipe\*\*\*\*

### ASSESSMENT TOOL FOR HEALTH-PROMOTION SPORTS PROGRAMMES

KEYWORDS: Assessment Tool, checklist, Healthy sporting events, Kappa coefficient, *V* Aiken.

ABSTRACT: The aim of this study was to develop a specific tool to assess the level of implementation of good practices in the organisation of sports programmes aimed at promoting and improving health. For this purpose, an ad hoc instrument based on a prior literature review called "Pilot Work" was designed. This document was submitted for review by a discussion group until an adequate level of consensus was obtained. Then a pilot test was conducted by two researchers, independently, selecting five programmes for good healthy sporting practices in Spain. The final version of the tool consisted of a total of 21 items, divided into two blocks: Common (4 items) and specific (17 items). The measuring instrument known as *EVESPRAH*, showed good validity ( $V$  Aiken  $\geq .96$ ) and reliability ( $K = .8122$ ) indices. This tool aims to guide planners of sports programmes for best practices in health promotion, active participation and quality, which can lead to an increase in adherence to physical activity in different populations.

Actualmente, las actividades físico-deportivas enfocadas a la mejora de la salud de las personas están en auge (García-González, Morales-Sánchez, Hernández-Mendo y Anguera, 2001; Sturm, 2004). El motivo principal son los numerosos beneficios que aporta la actividad física en términos de salud y calidad de vida (Pedrero et al., 2012), y la reducción del riesgo de padecer afecciones cardíacas, diabetes y algunos tipos de cáncer (Cenarruzabeitia, Hernández y Martínez-González, 2003). Además, la participación en este tipo de actividades previene el sobrepeso y mejora la salud de músculos y huesos, incluyendo beneficios psicológicos (Márquez, Rodríguez-Ordax y De Abajo, 2006). En los últimos años, una de las estrategias para fomentar este tipo de actividades han sido los programas o eventos deportivos basados en recomendaciones y buenas prácticas en materia de salud (McCartney et al., 2010).

Los gastos de la mayor parte de estos eventos o programas deportivos son asumidos por empresas e instituciones, debido a la gran repercusión social que les aportan (Sáez-Rodríguez, 2008; Walo, Bull y Breen, 1996). Estos programas no sólo tienen una finalidad lúdica y saludable, sino que también realizan una gran labor social enfocada a diferentes colectivos. Sin embargo, muchos de estas programas no aplican ningún tipo de herramienta de evaluación sobre la calidad o nivel de aplicación de buenas prácticas (Morales-Sánchez, Hernández-Mendo y Blanco, 2005). Para ello, se necesitan una serie de recursos donde las organizaciones deportivas juegan un papel muy importante como planificador, inversor y gestor (Corti, Donovan y Holman, 1996; Hallal, Victoria, Azevedo y Wells, 2006; Steffen et al., 2006).

El término calidad en el contexto de un servicio ha sido definido como la diferencia que existe entre lo que se espera del servicio y lo que el consumidor percibe del mismo; algo crucial para la fidelidad de un programa o evento deportivo (MacKay y Crompton, 1988). Recientes investigaciones se han unido a este interés en identificar cuáles son los aspectos clave que garantizan la calidad en dichos eventos o programas de actividad física (Kelley y Turley, 2001). Los estudios que se han llevado a cabo se centran, en su mayoría, en valorar la calidad por parte del espectador y/o del usuario a través de escalas psicométricas (Calabuig, Burillo, Crespo, Mundina y Gallardo, 2010; Hightower, Brady y Baker, 2002; Shonk y Chelladurai, 2008).

Posiblemente, el estudio pionero fue el desarrollado por McDonald, Sutton y Milne (1995), en el que se analizó la calidad percibida y la satisfacción con los programas y eventos deportivos a través de la adaptación de la escala *Servqual*. Esta escala es de las más utilizadas para la investigación específica en eventos deportivos (Bigne, Martínez, Miquel y Andreu, 2003). Otros estudios crearon sus propias herramientas con el fin de adaptarlas a las características de cada evento. Así, Wakefield, Blodgett y Sloan (1996) llevaron a cabo una investigación para mejorar la satisfacción del espectador a través de la herramienta *Sportscape* valorando tres dimensiones: los signos, símbolos y artefactos, la distribución espacial y funcional y las condiciones ambientales.

Los estudios clásicos incluyen elementos como la calidad del servicio básico, la calidad del servicio periférico, la satisfacción percibida y su relación con futuras intenciones (Tsuiji, Bennett y Zhang, 2007), la calidad percibida (Calabuig y Crespo, 2009;

Correspondencia: José Luis Felipe, Facultad de Ciencias de la Actividad Física y el Deporte. Universidad Europea, Ed. Juan Mayorga. C/Tajo, s/n, 28670 Villaviciosa de Odón (Madrid). Email: joseluis.felipe@universidadeuropea.es

\* Grupo IGOID. Universidad de Castilla-La Mancha.

\*\* Universidad Católica San Antonio de Murcia (UCAM).

\*\*\* Universidad de Valencia.

\*\*\*\* Universidad Europea.

Fecha de recepción: 7 de Enero de 2015. Fecha de aceptación: 19 de Diciembre 2015.

Calabuig, Mundina, y Crespo 2010; Hernández-Mendo, 2001; Morales-Sánchez, Hernández-Mendo y Blanco, 2009; Morales-Sánchez y Gálvez-Ruiz, 2011; A. Nuviola, Grao-Cruces, Pérez-Turpin y Nuviola, 2012; Theodorakis, Alexandris, Tsigilis, y Karvounis, 2012; Yoshida y James, 2011), o el valor percibido (Calabuig, Burillo et al., 2010). Por otro lado, García-González et al. (2011) se han centrado en la calidad y permanencia del voluntariado deportivo. Mientras, Gallardo, García-Tascón y Burillo (2008) investigaron sobre la eficiencia de la gestión de las instalaciones deportivas para ofrecer servicios de alta calidad a sus usuarios a través de los recursos materiales, las necesidades de los usuarios y la gestión de los recursos humanos. Dichos estudios se basan en encuestas de satisfacción y en la calidad percibida por los distintos agentes.

No obstante, salvo algunas pocas recomendaciones, actualmente no existe ninguna herramienta que permita evaluar la calidad de los programas a través del análisis del nivel de implantación de buenas prácticas en materia de salud, estandarizando los resultados en una misma magnitud. Los programas de buenas prácticas deportivas saludables se definen como los proyectos y/o actividades que a través de su estructura, recursos, planificación y valoración, contribuyen a crear hábitos saludables para la mejora de la calidad de vida del ciudadano (Australian Qualifications Framework Council, 2011). Por ello, el objetivo del presente estudio fue diseñar y validar una herramienta que permita valorar, de forma objetiva, la calidad y buenas prácticas de los programas deportivos saludables. Dicha herramienta debería recoger una serie de indicadores que analicen el nivel de implantación de buenas prácticas en la promoción de la salud, desde un punto de vista técnico y no subjetivo, por parte de los organizadores de programas/eventos/actividades deportivas.

## Método

### Participantes

Para poder validar la herramienta de evaluación de programas deportivos para la promoción de la salud se seleccionaron 5 programas de este tipo organizados en España, todos con características similares. Los 5 eran programas deportivos regionales, orientados al fomento del deporte para todos, independientemente de la edad, cuyo objetivo general fuese: contribuir al fomento de la práctica deportiva de la población participante. Todos ellos estuvieron subvencionados con financiación pública, con una duración comprendida entre 4 y 6 semanas (además presentaron al menos 3 sesiones por semana, con una duración superior a los 50').

### Instrumento

El primer paso a la hora de diseñar la herramienta fue elaborar un documento atendiendo a criterios básicos para el desarrollo de buenas prácticas basada en la revisión bibliográfica previa sobre la temática de estudio (Pena, 2009; Ysa, 2007). Este primer documento, llamado "Piloto de Trabajo" quedó compuesto por 48 indicadores agrupados en cuatro potenciales bloques de buenas prácticas relacionados con diferentes partes específicas en la organización de un programa deportivo. Para continuar el proceso de validación de la herramienta, a este Piloto de Trabajo se le aplicó una valoración por un Grupo de Discusión, a través de un panel de expertos en la materia, siguiendo la metodología propuesta en diferentes investigaciones (Anguera et al., 1995;

Creswell, 2003). Se utilizó este método por ser considerado el más apropiado para la indagación de un nuevo producto/servicio, o para el rediseño de uno ya existente (Queraltó, 2001; Yauch y Steudel, 2003). El Grupo de Discusión quedó compuesto por 20 personas, nueve hombres (45%) y 11 mujeres (55%) de diferentes países europeos: Reino Unido ( $n = 5$ ; 25%), Italia ( $n = 4$ ; 20%), Chipre ( $n = 2$ ; 10%), Hungría ( $n = 2$ ; 10%) y España ( $n = 7$ ; 35%). Todos ellos expertos en diferentes campos de especialización relacionados con la actividad física y la salud, arquitectura, poblaciones especiales y gestión deportiva. Los expertos participantes debían tener un mínimo de 10 años de experiencia en su ámbito. Todos ellos participaron de forma voluntaria y desinteresada. Asimismo, el 60% de los expertos formaban parte en el desarrollo de un proyecto europeo de programas deportivos orientados a la promoción de la salud, y al 40% restante se les envió un email informándoles del estudio junto con el documento Piloto de Trabajo elaborado.

Siguiendo las sugerencias de Merton (1987), el grupo se dividió en dos subgrupos de forma equilibrada en cuanto a género y especialidad profesional, para obtener percepciones y opiniones similares, así como evitar actitudes y sentimientos de prejuicios, hostilidad, etc., que pudieran influir entre los participantes. Así, se formaron dos subgrupos con 10 expertos cada uno, atendiendo a las investigaciones previas que sugieren que los grupos deben estar formados entre cinco y doce personas máximo, para favorecer la mayor oportunidad de participar (Krueger y Cassey, 2009). Las características del Grupo de Discusión y la división realizada puede verse en la Tabla 1.

El Grupo de Discusión se reunió en dos ocasiones. Al inicio de cada sesión se entregó una copia del guión de la misma a cada participante, se informó que la reunión sería grabada en audio para el correcto análisis posterior, y se indicó que sus comentarios serían tratados de forma confidencial y sin identificación personal. Además, en cada subgrupo y sesión se incluyó un investigador que sólo asistió como moderador, siendo especialistas en la temática de estudio. Los objetivos de las sesiones del Grupo de Discusión fueron identificar variables que inciden en la calidad de los programas deportivos orientados hacia la salud y corregir el reparto de las variables identificadas en bloques relacionados con las partes que forman parte de la organización de un programa deportivo. Las sesiones se desarrollaron en inglés.

La sesión 1 (por grupos) estuvo dividida en tres partes: (i) Presentación de los bloques y variables de análisis. Se expusieron los objetivos de la sesión y cómo se iba a desarrollar. Se insistió que se debía respetar el turno de palabra de cada intervención. (ii) Desarrollo y debate sobre las preguntas planificadas y aspectos clave para la presentación de los indicadores. Finalmente, cada indicador o variable se construyó a través de un ítem con formato de pregunta tipo checklist, cuyas respuestas se basaban en una rúbrica, específica y diferente para cada ítem, en la que existían cinco posibles respuestas correspondientes a 5 posibles prácticas (0 a 4), valorando el nivel de implantación de las mismas (el valor 0 representaba la peor práctica y el 5 la mejor práctica). Los miembros del Grupo de Discusión valoraron en este apartado la validez de contenido de la herramienta. Para ello, evaluaron cada ítem (incluyendo pregunta y las posibles respuestas de la rúbrica) con una escala Likert de cinco puntos, siendo 0 "ítem totalmente aceptable" y 4 "ítem totalmente rechazable". (iii) Conformación definitiva de los indicadores y sus respectivos bloques.

<b>Composición del Grupo</b>	<b>Grupo Discusión 1</b>	<b>Grupo Discusión 2</b>
Equipo Investigación	Un Moderador/Investigador	Un Moderador/Investigador
Expertos	Cuatro Gestores Deportivos Dos Expertos en Tercera Edad Un Experto Infancia y Juventud Dos Expertos en Tercera Edad Un Experto en Discapacidad Un Profesor Universidad experto en Actividad Física y Salud Un Arquitecto	Un Arquitecto Dos Expertos en Infancia y Juventud Dos Gestores Deportivos Un Profesor Universidad experto en Gestión Deportiva
Género	Cuatro Hombres Seis Mujeres	Cinco Hombres Cinco Mujeres
Nacionalidad	Dos Italia Uno Hungría Cuatro España Uno Chipre Dos Reino Unido	Dos Italia Uno Hungría Tres España Uno Chipre Tres Reino Unido
Perfil Profesional	Gestores de entidades deportivas (ámbito público y privado) Arquitectos especializados en el diseño de instalaciones deportivas Expertos en accesibilidad y personas con dependencia Expertos en juventud, animación, ocio y tiempo libre Profesores universitarios doctores especialistas en Actividad Física y Salud y en Gestión Deportiva	
Sesión 1	1 h 50' (10.00 - 11.50 h)	1 h 50' (10.00 - 11.50 h)
Sesión 2 (Conjunta)	1 h 15'	(12.30-13.45 h)

Tabla 1. Composición del Grupo de Discusión, División en Subgrupos, Perfil de los Participantes y Descripción de la Sesión.

Tras un descanso, de un día, donde el equipo investigador reformuló los ítems señalados por el Grupo de Discusión, se realizó la sesión dos (conjunta). Ésta estuvo dividida en dos partes: (i) Consenso sobre los indicadores imprescindibles y la puntuación que todos los programas debían cumplir para ser programas de buenas prácticas saludables. Se acordó emitir una calificación cualitativa donde era necesario establecer una puntuación mínima para aprobar, y a partir de ahí establecer

franjas de puntuación, dándole un nombre a cada una de ellas. (ii) Volver a evaluar cada ítem de nuevo con la escala tipo Likert utilizada en la sesión uno. Finalmente la herramienta quedó compuesta por 21 ítems, agrupados en una parte común y otra específica, esta última dividida a su vez en cuatro bloques relacionados con las partes que componen la organización de un programa deportivo (Tabla 2).

Partes y bloques	Descripción	
Parte común		
1. Transversalidad	La relación existente con otras áreas o materias	
2. Innovación	La capacidad del programa para sorprender a los usuarios	
3. Reconocimiento	La repercusión social del programa	
4. Transferibilidad	Su aplicabilidad a otras situaciones o circunstancias	
Parte específico		
2.1. Bloque Organizativo		
5. Planificación	La necesidad de realizar algunas acciones, que contribuyan al buen desarrollo de los programas de actividad física, deportivos; al buen funcionamiento y coordinación de los recursos humanos y materiales	
6. Flexibilidad		
7. Condiciones de seguridad		
8. Responsabilidad		
2.2. Bloque Programa		
9. Memoria del programa	La capacidad organizativa interna del programa de buenas prácticas deportivas saludables, que hacen que un programa deportivo esté bien coordinado, estructurado y organizado	
10. Evaluación		
11. Población diana		
12. Cualificación del profesorado		
2.3. Bloque Dimensión personal		
13. Calidad de vida y salud	El fomento de la práctica deportiva en la sociedad, la participación activa del usuario, el aumento en la adherencia a la actividad física y fomentar el sentimiento de pertenencia a un grupo	
14. Recomendaciones de actividad física		
15. Intensidad		
16. Tiempo de práctica		
17. Revisión médica previa		
18. Satisfacción del usuario/a		
2.4 Bloque Repercusión del programa		
19. Difusión		Establecer si los efectos del programa en la sociedad son significativos y perdurables en el tiempo y corregir los aspectos negativos para potenciar los efectos positivos de la práctica
20. Medidas medioambientales		
21. Participación		

Nota: La herramienta completa puede consultarse en: <http://www.webcitation.org/6Zlh29SLd>

Tabla 2. Diseño Definitivo de la Herramienta *EVESPRAH*.

La suma de la puntuación total máxima será de 105 puntos. Los criterios que se definieron poder catalogar a los programas como programas de buenas prácticas deportivas saludables fueron los siguientes: (i) Alcanzar un mínimo de 53 puntos (50 % de los puntos). (ii) En el bloque común es imprescindible conseguir como mínimo un tres en cada uno de los cuatro indicadores (Innovación, reconocimiento, transversalidad y transferibilidad).

#### Procedimiento

Para el análisis de validez de contenido a través de las reuniones con el Grupo de Discusión se calcularon los valores promedios de las puntuaciones obtenidas por la escala likert con la que cada experto valoró (Bulger y Housner, 2007). Igualmente, se utilizó la prueba *V* de Aiken (Merino y Livia, 2009; Penfield y Giacobbi, 2004), siguiendo la metodología descrita por Gómez, Sainz de Baranda, Ortega, Contreras y Olmedilla (2014).

Tras ello, se realizó una prueba piloto de la herramienta *EVESPRAH* tras la validación del contenido por el Grupo de Discusión, con el objetivo de confirmar la validez y fiabilidad de la herramienta. Para ello, se evaluaron los cinco programas deportivos seleccionados. Durante el proceso, el cuestionario fue traducido del inglés al castellano a través del procedimiento de traducción inversa (Ramada-Rodilla, Serra-Pujadas y Delclós-Clanchet, 2013). Cada programa fue evaluado por dos investigadores de manera independiente, contrastando posteriormente los resultados. Con los datos obtenidos, se realizó el análisis de la fiabilidad de la herramienta utilizando la prueba de Coeficiente de Kappa (Cohen, 1960; Landis y Koch, 1977), siguiendo el procedimiento descrito por Cerdá y Villarroel (2008). Para todo el análisis estadístico se utilizó el paquete estadístico SPSS v.20.

## Resultados

Una vez creado el Piloto de Trabajo, el Grupo de Expertos evaluó todos los ítems así como la rúbrica de evaluación de cada uno de ellos. Una vez recogidas todas las opiniones y llegados a una conclusión final, el equipo investigador diseñó el segundo Piloto de Trabajo, siguiendo la propuesta de Bulger y Housner (2007). Después se realizó la primera valoración del Grupo de Expertos, obteniendo una media de sus valoraciones con una escala Likert de 1 a 5, calculando la *V* de Aiken para cada uno de los ítems propuestos (Tabla 3). Para esta primera versión se decidió modificar todos aquellos ítems con valores inferiores a

0.90, y aceptar todos los que presentaran un valor superior a 0,95. En la segunda reunión del Grupo de Expertos se volvió a realizar la evaluación con los ítems modificados. El Grupo de Expertos volvió a emitir su veredicto y se volvieron a modificar todos los ítems con un valor inferior a 0.95. Tras esta reunión final, el equipo de investigación modificó nuevamente los ítems. Una vez modificados se envió vía correo electrónico el cuestionario definitivo a cada uno de los miembros del Grupo de Expertos, que telemáticamente volvió a emitir su juicio, llegando así a la versión definitiva del cuestionario, aceptando todas las variables al presentar valores comprendidos entre 0.95-1.00 (la media de valoración fue superior a 4.80 en todos los ítems valorados).

Ítems	Primera versión		Segunda Versión		Versión Final	
	<i>M</i>	<i>V</i> de Aiken	<i>M</i>	<i>V</i> de Aiken	<i>M</i>	<i>V</i> de Aiken
Parte común						
Transversalidad	3.46	.69	4.62	.92	5.00	1.00
Innovación	3.32	.66	4.13	.83	4.95	.99
Reconocimiento	3.59	.71	4.02	.82	4.87	.97
Transferibilidad	-	-	4.64	.92	4.90	.99
Parte específica						
Organizativo						
Planificación	3.40	.68	4.47	.91	4.99	.99
Flexibilidad	3.93	.78	4.37	.90	5.00	1.00
Condiciones de seguridad	4.05	.80	4.84	.96	4.97	.99
Responsabilidad	-	-	4.22	.84	5.00	1.00
Programa						
Memoria del programa	3.32	.66	4.46	.91	4.87	.97
Evaluación	3.46	.69	4.52	.92	4.88	.97
Población diana	4.23	.89	4.50	.91	5.00	1.00
Cualificación del profesorado	4.45	.91	4.66	.92	5.00	1.00
Dimensión personal						
Calidad de vida y salud	4.00	.80	4.69	.92	5.00	1.00
Recomendación de actividad física	3.98	.80	4.48	.91	4.89	.97
Intensidad	3.72	.74	4.46	.91	4.90	.98
Tiempo de práctica	3.23	.64	4.52	.92	5.00	1.00
Revisión médica previa	-	-	4.33	.90	4.85	.97
Satisfacción del usuario	4.21	.89	4.50	.91	5.00	1.00
Repercusión del programa						
Difusión	3.62	.73	4.48	.91	5.00	1.00
Medidas medioambientales	-	-	4.63	.92	4.82	.97
Participación	3.30	.66	4.59	.92	4.80	.96
Manera de puntuar	3.68	.73	4.33	.90	4.90	.98
Grado de dificultad	4.40	.90	4.86	.97	5.00	1.00

Tabla 3. Media y *V* de Aiken de las Valoraciones del Grupo de Expertos.

Seguidamente se realizó el análisis de la concordancia inter-observador (fiabilidad) mediante el cálculo del Coeficiente de Kappa (Tablas 4 y 5). Dos investigadores realizaron de manera independiente la evaluación de 5 programas deportivos celebrados en España. Cada investigador realizó un total de 105 evaluaciones con la herramienta EVESPAH (21 ítems x cinco

eventos). En la Tabla 4 se muestran las concordancias en las valoraciones de cada uno de los ítems (celdas a, b, c, d) así como sus discordancias (resto de celdas). A partir de estos datos se calcula el coeficiente de concordancia, al que se llamó de manera arbitraria  $\alpha$ , que corresponde al ratio entre el sumatorio de las observaciones concordantes y el total de observaciones.

INVESTIGADOR 1	INVESTIGADOR 2						TOTAL	$^1\alpha$
	0	1	2	3	4			
0	6 (a)	1	1	0	0	8	.8667	
1	2	22 (b)	0	1	1	26		
2	2	1	18 (c)	2	0	23		
3	0	0	0	40(d)	2	42		
4	0	0	0	1	5 (e)	6		
TOTAL	10	24	19	44	8	105		

$$\alpha = \frac{(\sum \text{concordancias observadas})}{\text{total de observaciones}}$$

Tabla 4. Valoración de cada uno de los programas de la prueba piloto del observador uno vs. observador dos.

A continuación, se calcularon todas las concordancias atribuibles al azar (a', b', c', d'). Partiendo de estos valores se

calculó el Coeficiente de Kappa ( $\kappa = 0.8122$ ) (Tabla 5).

INVESTIGADOR 1	INVESTIGADOR 2						TOTAL	$^1K$
	0	1	2	3	4			
0	2.29 (a')	-	-	-	-	8	.8122	
1	-	5.94 (b')	-	-	-	26		
2	-	-	4.16 (c')	-	-	23		
3	-	-	-	17.6 (d')	-	42		
4	-	-	-	-	0.46 (e')	6		
TOTAL	10	24	19	44	8	105		

$$\kappa = \frac{[(\sum \text{concordancias observadas}) - (\sum \text{concordancias atribuidas al azar})]}{[(\text{total observaciones}) - (\sum \text{concordancias atribuidas al azar})]}$$

(a') = (8x10/105); (b') = (26x24/105); (c') = (23x19/105); (d') = (44x42/105); (e') = (6x8/105)

Tabla 5. Valores Teóricos para cada Celda de la Diagonal Principal.

## Discusión

El objetivo del estudio fue crear una herramienta para valorar objetivamente la calidad del servicio de los programas y eventos deportivos saludables. Cada vez es más frecuente que personas de diferente clase social, participen en eventos deportivos y programas deportivos respaldados por la Administración (Morales, Hernández-Mendo y Blanco, 2005). Debido a la

situación económica y a la legislación vigente, éstas confían la gestión de los servicios del evento favoreciendo la creación de programas sin una metodología evaluativa necesaria (Anguera et al., 1995; Chacón, Anguera y López-Ruiz, 2000).

Es necesario poner a disposición de los técnicos o los directores de estos programas una herramienta que evalúe el nivel de calidad técnica de los programas deportivos que organizan, atendiendo al uso y alcance de buenas prácticas, puesto que de la



calidad de los mismos dependerá la satisfacción y resultados de los usuarios, y por ende, el nivel de práctica deportiva y el fomento de hábitos de vida saludables (Lloréns y Fuentes, 2000; Gallardo et al., 2008). La herramienta definitiva muestra valores promedios de las evaluaciones del Grupo de Expertos muy altas en todos los ítems y unos valores de  $V$  de Aiken muy adecuados. Estos resultados ponen de manifiesto la validez de contenido de la herramienta, ya que presenta valores promedio que indican la “alta aplicabilidad” de la misma (Bulger y Housner, 2007), así como valores de la  $V$  de Aiken óptimos (Pendfield y Giacobbi, 2004). Además, el Coeficiente de Kappa obtenido tras las pruebas piloto realizadas por dos investigadores de manera independiente, refleja la fuerza de concordancia “casi perfecta” obtenida con la herramienta *EVESPRAH* (Landis y Koch, 1977). Finalmente, cabe señalar como principal limitación del estudio el escaso número de programas deportivos analizados. Si bien con este

número de programas se obtienen unos valores adecuados de validez y fiabilidad de la herramienta, las investigaciones futuras deben pasar por evaluar un mayor número de programas deportivos de diferentes países europeos.

## Conclusiones

La herramienta *EVESPRAH* presenta unos índices de validez y fiabilidad adecuados. Por tanto, puede ser utilizada para puntuar eventos deportivos, catalogándolos como programas de buenas prácticas deportivas saludables y poder realizar comparaciones con otros programas similares, para realizar *benchmarking*, sin necesidad de recurrir a escalas con un mayor coste de aplicación o a cuestionarios de carácter científico que requieran de un análisis estadístico complejo.

### HERRAMIENTA DE EVALUACIÓN DE LOS PROGRAMAS DEPORTIVOS PARA LA PROMOCIÓN DE LA SALUD

**PALABRAS CLAVE:** Herramienta de evaluación, Checklist, Programas deportivos saludables, Coeficiente de Kappa,  $V$  de Aiken.

**RESUMEN:** El objetivo de este estudio fue elaborar una herramienta específica para evaluar el nivel de implantación de buenas prácticas en la organización de programas deportivos orientados al fomento y mejora de la salud. Para ello, se diseñó un instrumento ad hoc partiendo de una revisión bibliográfica previa al que se llamó “Piloto de Trabajo”. Este documento fue sometido a la revisión de un grupo de discusión hasta obtener un nivel de consenso adecuado. Seguidamente, se realizó una prueba piloto por dos investigadores de manera independiente, seleccionando 5 programas de buenas prácticas deportivas saludables en España. La versión final de la herramienta quedó compuesta por un total de 21 ítems, divididos en dos bloques: común (4 ítems) y específico (17 ítems). El instrumento de medida, denominado *EVESBEPRAH*, mostró buenos índices de validez ( $V$  de Aiken  $\geq .96$ ) y fiabilidad ( $K = .8122$ ). Esta herramienta pretende orientar a los planificadores de programas deportivos hacia las buenas prácticas en la promoción de la salud, la participación activa y la calidad, lo que puede implicar un aumento en la adherencia a la actividad física en diversas poblaciones.

### FERRAMENTA DE AVALIAÇÃO DE PROGRAMAS DESPORTIVOS PARA A PROMOÇÃO DA SAÚDE

**PALAVRAS-CHAVE:** Ferramenta de avaliação; checklist; programas desportivos saudáveis; Coeficiente de Kappa;  $V$  de Aiken.

**RESUMO:** O objectivo deste estudo foi elaborar uma ferramenta específica para avaliar o nível de implantação de boas práticas na organização de programas desportivos orientados para o fomento e melhoria da saúde. Para tal, delineou-se um instrumento ad hoc partindo de uma revisão bibliográfica prévia à que se chamou “Piloto de Trabalho”. Este documento foi submetido a uma revisão de um grupo de discussão até ser obtido um nível de consenso adequado. Seguidamente, realizou-se uma prova piloto por dois investigadores independentes, seleccionando 5 programas de boas práticas desportivas saudáveis em Espanha. A versão final da ferramenta ficou composta por um total de 21 itens, divididos em dois blocos: comum (4 itens) e específico (17 itens). O instrumento de medida, denominado *EVESBEPRAH*, mostrou bons índices de validade ( $V$  de Aiken  $\geq .96$ ) e fidelidade ( $K = .8122$ ). Esta ferramenta pretende orientar a planificação de programas desportivos visando as boas práticas na promoção da saúde, a participação activa e a qualidade, o que pode implicar um aumento na adesão à actividade física em diversas populações.

## Referencias

- Anguera, M. T., Arnau, J., Ato, M., Martínez, R., Pascual, J. y Vallejo, G. (1995). *Métodos de investigación en psicología*. Madrid: Síntesis.
- Australian Qualifications Framework Council, (2011). *Australian Qualifications Framework First Edition 2011*. Adelaide: Australian Qualifications Framework Council.
- Bigne, J. E., Martínez, C., Miquel, M. J. y Andreu, L. (2003). Servqual reliability and validity in travel agencies. *Annals of Tourism Research*, 30, 258-262.
- Bulger, S. M. y Housner, L. D. (2007). Modified Delphi investigation of exercise science in physical education teacher education. *Journal of Teaching in Physical Education*, 26, 57-80.
- Calabuig, F., Burillo, P., Crespo, J., Mundina, J. J. y Gallardo, L. (2010). Satisfacción, calidad y valor percibido en espectadores de atletismo. *Revista Internacional de Medicina y Ciencias de la Actividad Física y el Deporte*, 10, 577-593.
- Calabuig, F. y Crespo, J. (2009). Uso del método Delphi para la elaboración de una medida de la calidad percibida de los espectadores de eventos deportivos. *Retos: nuevas tendencias en Educación Física, Deporte y Recreación*, 15, 18-24.
- Calabuig, F., Mundina, J. y Crespo, J. (2010). Eventqual: una medida de la calidad percibida por los espectadores de eventos deportivos. *Retos: Nuevas tendencias en Educación Física, Deporte y Recreación*, 18, 66-70.
- Cenarruzabeitia, J. J., Hernández, J. M. y Martínez-González, M. A. (2003). Beneficios de la actividad física y riesgos del sedentarismo. *Medicina Clínica*, 121, 665-672.
- Cerdá, J. y Villarroel, L. (2008). Evaluación de la concordancia inter-observador en investigación pediátrica: Coeficiente de Kappa. *Revista Chilena de Pediatría*, 79, 54-58.
- Chacón, S., Anguera, M. T. y López-Ruiz, J. (2000). Diseños de evaluación de programas: bases metodológicas. *Psicothema*, 12(Suplemento), 127-131.
- Cohen, J. A. (1960). Coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 20, 37-46.
- Corti, B., Donovan, R. J. y Holman, C. D. J. (1996). Factors influencing the use of physical activity facilities: results from qualitative research. *Health Promotion Journal of Australia*, 6, 16-21.

- Creswell, J. W. (2003). *Research design: Qualitative, quantitative and mixed methods approaches*. Thousand Oaks, CA: Sage.
- Gallardo, L., García-Tascón, M. y Burillo, P. (2008). New sports management software: A needs analysis by a panel of Spanish experts. *International Journal of Information Management*, 28, 235-245.
- García-González, R., Morales-Sánchez, V., Hernández-Mendo, A. y Chica-Merino, E. (2011). Una herramienta para evaluar la calidad y la permanencia en el voluntariado deportivo. *Cuadernos de Psicología del Deporte*, 11, 171-178.
- Gómez, P., Sainz de Baranda, P., Ortega, E., Contreras, O. y Olmedilla, A. (2014). Diseño y validación de un cuestionario sobre la percepción del deportista respecto a su reincorporación al entrenamiento tras una lesión. *Revista de Psicología del Deporte*, 23, 479-487.
- Hallal, P. C., Victora, C. G., Azevedo, M. R. y Wells, J. C. (2006). Adolescent physical activity and health: a systematic review. *Sports Medicine*, 36, 1019-1030.
- Hernández-Mendo, A. (2001). Un cuestionario para evaluar la calidad en programas de actividad física. *Revista de Psicología del Deporte*, 10, 179-196.
- Hernández-Mendo, A. y Anguera, M.T. (2001). Análisis psicosocial de los programas de actividad física: Evaluación de la temporalidad. *Psicothema*, 13, 263-270.
- Hightower, R., Brady, M. K. y Baker, T. L. (2002). Investigating the role of the physical environment in hedonic service consumption: an exploratory study of sporting events. *Journal of Business Research*, 55, 697-707.
- Kelley, S. W. y Turley, L. W. (2001). Consumer perceptions of service quality attributes at sporting events. *Journal of Business Research*, 54, 161-166.
- Krueger, R. A. y Cassey, M. A. (2009). *Focus groups: a practical guide for applied research*. Thousand Oaks, CA: Sage.
- Landis, J. y Koch, G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33, 159-74.
- Lloréns, F. J. y Fuentes, M. M. (2000). *Calidad Total: Fundamentos e Implantación*. Madrid: Pirámide.
- MacKay, K. J. y Crompton, J. L. (1988). A conceptual model of consumer evaluation of recreation service quality. *Leisure Studies*, 7, 41-49.
- Márquez, S., Rodríguez-Ordax, J., y De Abajo, S. (2006). Sedentarismo y salud: efectos beneficiosos de la actividad física. *Apuntes: Educación Física y Deporte*, 83, 12-24.
- McCartney, G., Thomas, S., Thomson, H., Scott, J., Hamilton, V., Hanlon, P., . . . Bond, L. (2010). The health and socioeconomic impacts of major multi-sport events: systematic review (1978-2008). *British Medical Journal*, 340, c2369. doi: <http://dx.doi.org/10.1136/bmj.c2369>.
- McDonald, M. A., Sutton, W. A. y Milne, G. R. (1995). TEAMQUAL™: Measuring service quality in professional team sports. *Sport Marketing Quarterly*, 4, 9-15.
- Merino, C., y Livia, J. (2009). Intervalos de confianza asimétricos para el índice la validez de contenido: un programa visual basic para la V de Aiken. *Anales de Psicología*, 25, 169-171.
- Merton, R. K. (1987). The focussed interview and focus groups: Continuities and discontinuities. *The Public Opinion Quarterly*, 51, 550-566.
- Morales-Sánchez, V., y Gálvez Ruiz, P. (2011). La percepción del usuario en la evaluación de la calidad de los servicios municipales deportivos. *Cuadernos de Psicología del Deporte*, 2, 147-154.
- Morales-Sánchez, V., Hernández-Mendo, A. y Blanco, A. (2005). Evaluación de la calidad en los programas de actividad física. *Psicothema*, 17, 311-317.
- Morales-Sánchez, V., Hernández-Mendo, A. y Blanco, A. (2009). Evaluación de la calidad en organizaciones deportivas: adaptación al modelo Servqual. *Revista de Psicología del Deporte*, 18, 137-150.
- Nuviala, A., Grao-Cruces, A., Pérez-Turpin, J. A. y Nuviala, R. (2012). Perceived service quality, perceived value and satisfaction in groups of users of sports organizations in Spain. *Kinesiology*, 44, 94-103.
- Pedrero, R., Gómez-Cabello, A., Delgado, S., Rodríguez-Llarena, S., Rodríguez-Marroyo, J. A., Cabanillas, E., . . . González-Gross, M. (2012). Physical fitness levels among independent non-institutionalized Spanish elderly: The elderly EXERNET multi-center study. *Archives of Gerontology and Geriatrics*, 55, 406-416.
- Pena, L. (2009). *Criteris per l'avaluació de les bones pràctiques en esports*. Barcelona: Seae, Planificació i Consultoria y Barcelona Provincial Council.
- Penfield, R. D. y Giacobbi, P. R. (2004). Applying a score confidence interval to Aiken's item content-relevance index. *Measurement in Physical Education and Exercise Science*, 8, 213-225.
- Queralto, P. (2001). *Investigación en técnicas de análisis cualitativo: el grupo de discusión*. Barcelona: Universitat Politècnica de Catalunya. Departament de Construccions Arquitectòniques I.
- Ramada-Rodilla, J. M., Serra-Pujadas, C. y Delclós-Clanchet, G. L. (2013). Adaptación cultural y validación de cuestionarios de salud: revisión y recomendaciones metodológicas. *Salud Pública de México*, 55, 57-66.
- Sáez-Rodríguez, G. (2008). El patrocinio deportivo y la imagen del deportista. *International Journal of Sports Law y Management*, 1, 3-10.
- Shonk, D. J. y Chelladurai, P. (2008). Service quality, satisfaction, and intent to return in event sport tourism. *Journal of Sport Management*, 5, 587-602.
- Steffen, L. M., Arnett, D. K., Blackburn, H., Shah, S., Armstrong, C., Luepker, R. V. y Jacobs, D. R. (2006). Population trends in leisure-time physical activity: Minnesota heart survey, 1980-2000. *Medicine and Science in Sports and Exercise*, 38, 1716-1723.
- Sturm, R. (2004). The economics of physical activity: societal trends and rationales for interventions. *American Journal of Preventive Medicine*, 27, 126-135.
- Theodorakis, N. D., Alexandris, K., Tsigilis, N. y Karvounis, S. (2012). Predicting spectators' behavioural intentions in professional football: The role of satisfaction and service quality. *Sport Management Review*, 16, 85-96.
- Tsuji, Y., Bennett, G. y Zhang, J. (2007). Consumer satisfaction with an action sports event. *Sport Marketing Quarterly*, 16, 199-208.
- Wakefield, K. L., Blodgett, J. G. y Sloan, H. J. (1996). Measurement and management of the sportscape. *Journal of Sport Management*, 10, 15-31.
- Walo, M., Bull, A., y Breen, H. (1996). Achieving economic benefits at local events: a case study of a local sports event. *Festival Management y Event Tourism*, 4, 95-106.
- Yauch, C. A., y Stuedel, H. J. (2003). Complementary use of qualitative and quantitative cultural assessment methods. *Organizational Research Methods*, 6, 465-479.
- Yoshida, M. y James, J. D. (2011). Service quality at sporting events: Is aesthetic quality a missing dimension? *Sport Management Review*, 14, 13-24.
- Ysa, T. (2007). Governance Forms in Urban Public-Private Partnerships. *International Public Management Journal*, 10, 35-57.