**DISEÑO Y VALIDACIÓN DE UN INSTRUMENTO DE OBSERVACIÓN DE CONDUCTAS SOBRE LAS CONDUCTAS DE CONVIVENCIA EN EDUCACIÓN FÍSICA**

**DESIGN AND VALIDATION OF AN OBSERVATION INSTRUMENT TO ASSESS STUDENT MISCONDUCT IN PHYSICAL EDUCATION**

Bernardino J. Sánchez-Alcaraz Martínez1, Alberto Gómez-Mármol2, Alfonso Valero-Valenzuela1, Rocío Esteban Luís1 y Sixto González-Villora3

1 Facultad de Ciencias del Deporte. Universidad de Murcia.

2 Facultad de Educación. Universidad de Murcia.

3 Facultad de Educación. Universidad de Castilla la Mancha.

Dirección para correspondencia

Bernardino J. Sánchez-Alcaraz Martínez 1

Facultad de Ciencias del Deporte.

Universidad de Murcia.

Campus de San Javier - C / Argentina s/n, San Javier, CP. 30720.

Email: [bjavier.sanchez@um.es](mailto:bjavier.sanchez@um.es)

**DISEÑO Y VALIDACIÓN DE UN INSTRUMENTO DE OBSERVACIÓN DE CONDUCTAS SOBRE LAS CONDUCTAS DE CONVIVENCIA EN EDUCACIÓN FÍSICA**

Bernardino J. Sánchez-Alcaraz Martínez, Alberto Gómez-Mármol, Alfonso Valero-Valenzuela, Rocío Esteban Luís y Sixto González-Villora

**Resumen**

El objetivo del presente trabajo fue diseñar y validar un instrumento de observación de conductas que alteran la convivencia en las clases de Educación Física, que permita definir, categorizar y medir cuantitativamente dichas conductas. Este proceso de validación se llevó a cabo en varias fases, contando con la participación de seis jueces expertos en el ámbito de la enseñanza de la Educación Física y cuatro observadores formados en el análisis de las conductas que alteran el clima de convivencia en las clases de Educación Física, que analizaron un total de diez clases de Educación Física, de sexto de Educación Primaria y de tercero de la Educación Secundaria de diez centros escolares diferentes, analizando la estructura del instrumento mediante procedimientos confirmatorios. El instrumento final estuvo formado por ocho conductas que alteran la convivencia en las clases de Educación Física, agrupadas en dos factores (conductas violentas y conductas de indisciplina y desinterés). Los resultados obtenidos en el proceso de validez, fiabilidad y consistencia interna indican que el instrumento diseñado es válido para registrar las conductas que alteran la convivencia en las clases de Educación Física, permitiendo su aplicación sin la presencia del observador en el aula, minimizando el efecto de la misma en el comportamiento del alumnado.

**Palabras clave:** población en edad escolar; instrumento de medida; Educación Física; evaluación de la educación.

**Abstract**

The aim of this study was to design and validate an instrument of observation that would allow us to measure and quantify student misconduct during classes of physical education. This process of validation was carried out in several phases. Six expert judges in physical education teaching and four observers trained to measure student misconduct in physical education lessons participated in the process, analysing ten physical education lessons of sixth and ninth grade from ten different schools, settings the structure of the instrument through confirmatory procedures. The final instrument consisted of eight items describing different kinds of student misconduct during physical education classes, grouped into two categories (violent behaviours and undisciplined and disinterested behaviours). The results obtained in the process of validation, reliability and internal consistency suggest that the instrument designed is optimally useful and reliable to measure student misconduct during physical education classes, allowing its application without the presence of the investigator in the classroom, thus minimizing its effect on the behaviour of the student body.

**Keywords:**school age population; measuring instruments; Physical Education; educational evaluation.

**Introducción**

Durante la adolescencia se experimentan, además de un rápido crecimiento físico, importantes modificaciones de índole conductual debidas a importantes cambios psico-sociales ligados al desarrollo (Escartí, Pascual, y Gutiérrez, 2005). Estos cambios psicosociales comienzan a acentuarse en inicio de la adolescencia, entre los 10 y 12 años, y suponen en muchas ocasiones una crisis de identidad y valores, así como un incremento de conductas que alteran la convivencia, lo que se traduce con demasiada frecuencia en un considerable incremento de conflictos familiares y/o educativos (Borràs, Palou, y Ponseti, 2001).

La convivencia educativa debe ser entendida como una construcción colectiva y dinámica compuesta a partir del conjunto de interrelaciones humanas entre los agentes de la comunidad educativa, cuyos límites se definen por derechos y deberes, cuya influencia trasciende del mero contexto escolar (García, 2010).

En los últimos años, se vive una crisis de valores y de convivencia en las aulas, traducido a un aumento del absentismo escolar, acompañado de un clima de relación social en los que las intimidaciones, el maltrato, la xenofobia, la violencia o las drogas empiezan a ser muy comunes en la mayoría de los centros escolares (Ruiz, Rodríguez, Martinek, Schilling, Durán, y Jiménez, 2006).

La participación de los estudiantes en comportamientos agresivos se ha asociado con resultados negativos, incluyendo el fracaso escolar, inadaptación social, rechazo, o comportamientos disruptivos y/o destructivos (Barnes, Smith, y Miller, 2014). En este sentido, las conductas disruptivas en el entorno escolar, como la hostilidad o la agresión, pueden poner en peligro el mantenimiento de ambientes de aprendizaje óptimos en las escuelas (Kuhn, Ebert, Gracey, Chapman, y Epstein, 2015).

Los cambios propios de la adolescencia pueden ser el origen de conflictos en el entorno educativo, y debido a la incidencia de conductas que alteran la convivencia, han proliferado en los últimos años estudios con programas de intervención orientados a la educación en materia de comportamiento socialmente adecuado. En estos programas se incluyen actividades físicas y/o deportivas, por ejemplo en fútbol de formación con la intención de inculcar valores de juego limpio (Palou, Borràs, Ponseti, Vidal, y Torregrosa, 2007), o bien con un desarrollo positivo para las comunidades de jóvenes que están en algún grado de marginación (Whitley, Forneris, y Barker, 2015).

En estudios recientes, también se ha evaluado experiencias novedosas desde la perspectiva del formador, poniendo el objetivo de análisis en los profesores o su metodología. Tanto en Educación Primaria (España), sirva de ejemplo el estudio cualitativo implementado el stop-motion como estrategia innovadora, basada en la animación y creación audiovisual, para la mejora de la convivencia. Tras esta intervención se constató la mejora del comportamiento de los estudiantes y un incremento en su motivación (Fernández del Río y Barreiras Arias, 2015). Como en Educación Secundaria (EE.UU.), con programas basados en el fomento de la responsabilidad, donde se encontró resultados positivos, pues se demostró que los profesores aumentaron el conocimiento de estrategias de enseñanza basadas en la responsabilidad, y percibieron un impacto positivo en los estudiantes (Hemphill, Templin, y Wright, 2015).

La mayor parte de los trabajos que estudian estos fenómenos se han realizado utilizando diferentes escalas y cuestionarios, o bien otro tipo de métodos cualitativos, como entrevistas (Urbina, Simón, y Echeita, 2011). Esta metodología, a pesar de su amplia aceptación y difusión, tiene sesgos propios que pueden afectar a la cuantificación concreta del número o la tipología de las conductas que alteran la convivencia en el entorno educativo. Una alternativa viable y que puede mejorar la calidad de la información obtenida en estos trabajos es la metodología observacional (Anguera, Blanco, Hernández, y Losada, 2011).

La observación permite el registro organizado y la cuantificación de las conductas perceptibles que se quieran analizar si se adecua un instrumento válido (Anguera y col., 2011). Dentro del contexto de convivencia escolar y, más concretamente, en el de conductas que alteran la convivencia en las clases de Educación Física, no existen muchos instrumentos que permitan registrar las conductas perceptibles.

En este sentido, una de las primeras hojas de observación, basada en el sistema de categorías, diseñada para evaluar conductas que alteran la convivencia y específica para el ámbito de la Educación Física, fue el System of Observation of Disciplinary Incidents, o bien SODI (Sistema de Observación de Incidentes de Disciplina), validado por Goyette, Dore y Dion (2000), que contiene 19 tipos de conductas inadecuadas, agrupadas en tres categorías según la intensidad (leve, grave o muy grave). Este sistema de categorías forma parte de un estudio centrado en estudiantes universitarios y futuros profesores, con la finalidad de descubrir y comprender las diferentes reacciones que los distintos profesores pueden aplicar en función del contexto en el que se manifiestan los comportamientos de los alumnos, por lo que no es válida para registrar cuantitativamente el número de conductas acaecidas en las clases.

Posteriormente, apareció otro sistema de categorías para observar las conductas antisociales (CAS) diseñada por Gras y Martínez (2002) que evaluaba la conducta de los escolares mientras realizaban actividad física. Este instrumento agrupa las conductas antisociales en seis categorías: agresión, quitar (robar), molestar, falta de control, forzar u obligar y la categoría otros. Se trata de una hoja de observación con ciertas limitaciones, como la exclusión en la categoría de “agresión” de las vertientes verbales y una definición poco operativa de la categoría “otros”. Por otro lado, este sistema de categorías presenta un diseño específico para la observación “in situ”, y que tal y como señala Anguera (2010), las tendencias actuales en investigación observacional tratan de evitar este tipo de observaciones para avanzar hacia la grabación de las conductas y su análisis informatizado, con el objeto de poder volver a observar una conducta que ya ha sucedido, o poder registrar varias conductas que acaecen simultáneamente.

En este sentido, Escartíy col. (2005) elaboraron un sistema de categorías para la evaluación de los comportamientos de responsabilidad personal y social, en la que, a través de la visualización del video de las clases de Educación Física, los observadores identifican de manera cuantitativa la aparición de las conductas de agredir, interrumpir, no colaborar y ayudar a lo largo de diferentes sesiones. Este sistema no ha sido validado científicamente, sino que presenta una concordancia consensuada, no profundizando en los análisis de calidad de dato (Hernández, Díaz y Morales, 2010).

Hernández, Díaz, y Morales (2010) expusieron una nueva herramienta observacional que contenía altos valores de fiabilidad y validez, específico para las clases de Educación Física, pero que registraba 22 conductas que alteran la convivencia, un número muy elevado, lo cual dificulta su aplicación, además de no incluir conductas tan importantes como las agresiones.

Por lo tanto, el limitado número de instrumentos de observación de conductas que alteran la convivencia en las clases de Educación Física, y los problemas existentes en los ya creados, tales como un elevado o bajo número de conductas a observar, registros cualitativos, técnicas de observación “in situ” o la no validación científica de alguna de ellas, ponen de manifiesto la necesidad del diseño de un nuevo instrumento que supla las carencias de las ya existentes.

De este modo, el objetivo principal de la investigación es diseñar y validar un instrumento de observación de conductas que alteran la convivencia en las clases de Educación Física, que permita definir y categorizar dichas conductas, facilitando su aplicación sin la presencia del observador en el aula y minimizando el efecto de la misma en el comportamiento del alumnado.

**Método**

*Muestra*

Para el desarrollo de la investigación se emplearon dos tipos de participantes. En primer lugar, se solicitó la colaboración de seis jueces expertos, todos ellos Maestros especialistas en Educación Física, Licenciados en Ciencias de la Actividad Física y el Deporte, docentes en activo y Doctores en Ciencias de la Actividad Física y el Deporte o Catedráticos de Educación Secundaria.

En segundo lugar, cuatro observadores, Licenciados en Ciencias de la Actividad Física y el Deporte, fueron formados en la valoración de las conductas que alteran el clima de convivencia, y se encargaron de visualizar el número de conductas y debatir sobre la tipología de las mismas en el proceso de validación del instrumento.

Para ello, se analizaron 10 clases de Educación Física de cursos de sexto de Educación Primaria y tercero de Educación Secundaria Obligatoria, de 10 Centros Escolares diferentes. La selección del muestreo de los centros fue por conveniencia, debido a la accesibilidad por parte de los investigadores. Los centros se encontraban en diferentes contextos socio-económicos. Las clases analizadas estaban compuestas por un número de entre 20 y 32 alumnos. En este caso, la selección de las clases dentro de cada centro educativo fue aleatorio dentro de los cursos elegidos en el estudio (sexto de Educación Primaria y tercero de Educación Secundaria).

*Procedimiento*

El procedimiento de la investigación estuvo formado por varias fases, en las que se realizaron diferentes versiones del instrumento definitivo. Durante este proceso, se siguió la línea de trabajo de los autores Cruz y col. (1996), en la validación de un instrumento de observación de conductas relacionadas con el fair play en fútbol.

*Fase 1. Elaboración de la primera versión del Instrumento de Observación de Conductas que Alteran la Convivencia en las clases de Educación Física (CACEF).*

En esta primera fase se elaboró una versión a partir del instrumento desarrollado por Kulinna, Cothran, y Recualos (2003), *Physical Education Classroom Management Instrument* (Instrumento de la Gestión del Aula en Educación Física), que abarca y evalúa todas la tipologías de comportamientos que alteran la convivencia en el área de Educación Física, tal y como se entiende actualmente este fenómeno en los centros. Esta adaptación transcultural realizada por Díaz y Esteban (2010) consta de 59 ítems que describen conductas inadecuadas, clasificadas en seis categorías, relacionadas con la agresividad, la falta de madurez, el desinterés académico, disruptivas, antisociales y de indisciplina.

Posteriormente, un grupo formado por seis jueces expertos buscó reducir el número de ítems de forma que fueran los más significativos por un lado y que se pudieran utilizar de una forma eficaz durante las sesiones de Educación Física, facilitando el entrenamiento de los observadores. El resultado fue una hoja de observación inicial con 12 ítems encaminados a determinar la frecuencia con la que se producían las conductas por parte de cualquiera de los alumnos en el desarrollo de la sesión (Tabla 1).

Tabla 1. Conductas utilizadas para confeccionar la primera versión del CACEF

|  |  |
| --- | --- |
| CONDUCTAS | |
| 1. Quejarse | 7. No seguir las indicaciones del profesor |
| 2. Hacer trampas, engañar en los juegos y deportes | 8. Interrumpir o molestar a los demás |
| 3. Reírse, burlarse de los compañeros | 9. Insultar, agredir verbalmente |
| 4. Agredir físicamente | 10. Realizar gestos obscenos |
| 5. Excluir a compañeros de juegos y deportes | 11. Amenazar a un compañero |
| 6. Recriminar, discutir irrespetuosamente | 12. Maltratar el material |

Fuente: elaboración propia.

*Fase 2. Ensayo, discusión y reestructuración de la primera versión del CACEF.*

En esta segunda fase, cuatro observadores analizaron en común dos sesiones de 60 minutos de una clase de Educación Física de sexto curso de Educación Primaria y de tercero de Educación Secundaria Obligatoria. De este modo, se realizó una primera prueba del instrumento de observación, en la que se analizaron y registraron las conductas, después de una puesta en común por parte de los observadores.

Cuando se finalizó con el análisis del video, y tras comentar los aspectos más relevantes sobre la facilidad o dificultad de registrar aquello observado en el monitor, se realizó una puesta en común para añadir, eliminar, modificar o agrupar las conductas de la plantilla inicial, elaborando una segunda versión.

- Conductas agrupadas: fueron la conducta “agredir verbalmente”, la conducta amenazar y la conducta “reírse y burlarse de los compañeros” agrupadas en una única conducta, denominada Agredir Verbalmente, ocupando el primer lugar en la hoja de registro.

- Conductas renombradas: se renombró la conducta “realizar gestos obscenos” pasando a llamarse “agredir gestualmente”, ocupando el tercer lugar en la hoja de registro.

- Conductas eliminadas: se eliminó la conducta “recriminar/discutir irrespetuosamente”, por considerar que podría suponer motivo de confusión con la conducta “agredir verbalmente” y la conducta “quejarse”.

*Fase 3. Ensayo, discusión y realización del instrumento definitivo.*

En esta ocasión, se utilizó la segunda versión de la plantilla para el análisis de otras dos sesiones de 50 minutos de otras dos clases de Educación Física. Posteriormente, en una sesión de laboratorio, se pusieron en común los registros obtenidos por los cuatro observadores con el fin de conocer las coincidencias y discrepancias entre ellos, así como la objetividad del instrumento de observación. Se determinó que únicamente en el caso de confusión y/o discrepancia entre los cuatro observadores, se volvía a repetir la acción hasta llegar a un acuerdo sobre cuál debía ser la conducta o categoría a la que la acción debía ser adscrita. Tras la discusión sobre los registros obtenidos, y la utilidad práctica de la plantilla, se decidió realizar algunas modificaciones en ésta:

- Conductas renombradas: se renombró la conducta “maltratar el material”, pasando a denominarse “maltratar el material o las instalaciones”, abarcando no exclusivamente el material de clase, sino también las instalaciones del centro.

En esta fase, una vez obtenido el modelo del instrumento final, se llevó a cabo la definición operacional de todas las conductas del instrumento de observación (Tabla 2), redefiniendo aquellas que habían planteado problemas a los observadores.

Tabla 2. Definición de conductas del instrumento de observación

|  |  |
| --- | --- |
| CONDUCTAS | DEFINICIÓN |
| Agredir físicamente | Acciones del alumnado en las que más allá de las situaciones deportivas en las que pueda existir contacto físico, los alumnos empujan, dan patadas, puñetazos, golpes, zancadillean, zarandean a otros compañeros. Por ejemplo: “*un alumno golpea a otro”* |
| Agredir verbalmente | Acciones del alumnado en las que se mofan de los compañeros o del profesor, los insultan, amenazan o faltan al respeto mediante la palabra, al igual que puede ser considerado como agresión verbal una exclusión a un alumno de un juego o deporte. |
| Agredir gestualmente | Situaciones en las que los alumnos se expresan con gestos inapropiados, ofensivos o con connotaciones sexuales. Por ejemplo: *“un alumno realiza un corte de mangas”.* También puede considerarse agresión gestual aquella conducta por la cual el alumno amenaza con pegar o maltratar a un compañero, aun sin llegar a hacerlo. Por ejemplo: *“un alumno alza el puño para pegar a otro, pero sin llegar al contacto”.* |
| Interrumpir o molestar a los demás | Acciones en las que los alumnos no dejan que la clase tenga su continuidad normal, ya sea durante la realización de las actividades o durante la explicación de las mismas. Se registrará como conducta de interrumpir o molestar a los demás, cada vez que el profesor interrumpa la clase pidiendo silencio o llamando la atención a algún alumno. Si el profesor pide silencio de forma general a toda la clase, se anotará como una única conducta de interrupción, pero si lo hace de forma específica, por ejemplo, a tres alumnos en concreto, se registrará como tres conductas de interrupción. Por ejemplo: *“un alumno comienza a hablar o a molestar a sus compañeros en el momento en el que el profesor explica una tarea”* |
| Maltratar el material o a las instalaciones | Se producen en los momentos en los que el alumno, utiliza de manera irresponsable el material, con excesiva fuerza, o realizando tareas con el mismo para el que no está diseñado y que puede provocar que se deteriore. Por ejemplo: *“el alumno da una patada a los conos de demarcación”.* El maltrato de las instalaciones también se considera una conducta dentro de maltrato de material*.* Por ejemplo: *“escupir en el suelo”.* |
| Hacer trampas, engañar en los juegos y deportes | Se produce cuando los alumnos, incumplen las normas de juegos y actividades con ánimo de conseguir la victoria o alguna ventaja para determinadas acciones o durante el desarrollo de un juego o actividad, o para dar ventaja a su equipo con estas acciones. Por ejemplo: “*los alumnos cuentan más goles en el partido de los que se han producido”* |
| Excluir a compañeros | Se produce en las situaciones en las que el alumnado, no deja participar, ignora, o minusvalora a algunos compañeros a la hora de realizar tareas propuestas por el profesor en las que se requiere la formación de equipos, grupos o parejas. Por ejemplo: “*un alumno queda sin equipo cuando se proponen para realizar una tarea”* |
| Quejarse | Se produce cuando el alumno expresa su descontento con la actividad o con los compañeros. Debe cumplir dos requisitos para que se considere una conducta que altera la convivencia: que no exista motivo objetivo para la queja y que se produzca de manera reiterada por el mismo alumno expresando su reclamación con salidas de tono, sin llegar a faltar al respeto, insultar o amenazar, en cuyo caso sería una conducta de agresión verbal. Por ejemplo: “*estoy cansado de hacer lo mismo siempre”* |
| No seguir las indicaciones del profesor | Situaciones en las que el alumnado ignora, realiza de otra manera, o no realiza los ejercicios y tareas, propuestos por el profesor, o no sigue las consignas propuestas para la sesión o tarea en concreto. Por ejemplo: “*un alumno comienza a tirar a puerta, cuando el profesor ha indicado que hay que realizar pases con el compañero*”. Por otro lado, se registrarán como “no seguir las indicaciones” tantas conductas realizadas por el alumno, pero si esta conducta es de forma continuada en un mismo ejercicio, sólo se anotará una vez. Por ejemplo: *“un alumno no participa en la actividad de fútbol, se registrará como una única conducta”.* Para considerarse conducta disruptiva, debe pasar un tiempo mínimo de 10 segundos continuado en los que el alumno no realice la actividad indicada por el profesor, ya que en ocasiones, algunos alumnos pueden pararse debido a que están cansados o están cambiando de un ejercicio a otro diferente. Es importante anotar esta conducta únicamente cuando el comportamiento del alumno no se corresponda a ninguna de las conductas anteriores. Por ejemplo: *“un alumno empuja a otro durante un juego, aunque el alumno no sigue la indicación del profesor de no empujar violentamente, la conducta es una agresión física”* |

Fuente: elaboración propia.

Finalmente, se añadieron apartados que mejoraban y ampliaban la información facilitada por el instrumento CACEF, tales como, secciones para conocer la fecha de observación de las clases de Educación Física, anotar el nombre del centro, el curso, el número de alumnos de la clase, el tiempo total de la sesión y una sección de observaciones (Ver Figura 1).

Figura 1. Versión definitiva del instrumento de observación de conductas que alteran la convivencia en clase de Educación Física

CENTRO: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ FECHA:\_\_\_\_/\_\_\_\_\_/\_\_\_\_\_

GRUPO: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ Nº ALUMNOS:\_\_\_\_\_\_\_\_ TIEMPO:\_\_\_\_\_min \_\_\_\_\_\_seg\_\_\_\_\_\_

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| CONDUCTA | MOMENTO DE APARACIÓN DE LA CONDUCTA  (Tiempo: min y seg) | | | | | | | | | | | | | | | | | | | | | | | | | | | FRECUEN-CIA TOTAL |
| 1. Agredir verbalmente |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 2. Agredir físicamente |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 3. Agredir gestualmente |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 4. Interrumpir o molestar a los demás |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 5. Maltratar el material o las instalaciones |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 6. Hacer trampas, engañar en los juegos y deportes |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 7. Quejarse |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 8. No seguir las indicaciones del profesor |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

OBSERVACIONES:

*Fase 4: Juicio de Expertos y análisis de los datos:*

El objetivo principal de esta fase fue adquirir la validez de las diferentes conductas que formaban el CACEF. Para llevar a cabo esta validación, el instrumento fue enviado a un panel de seis jueces expertos, adjuntando una carta junto al mismo, justificando su colaboración y el conocimiento del objeto de estudio. Dichos jueces observaron el grado de representatividad y calidad técnica de todas las conductas que forman el instrumento, realizando importantes aportaciones y valoraciones. Las aportaciones de los jueces expertos y los datos recogidos en la observación de las clases de Educación Física fueron analizados en el proceso de validación del instrumento, como se refleja en el apartado de resultados.

**Resultados**

*Validez del instrumento CACEF mediante jueces expertos*

Para considerar si las conductas son adecuadas para realizar las mediciones, los jueces expertos realizaron una revisión cualitativa de la hoja de observación mediante el Método Delphi (Landeta, 1999), del mismo modo que evaluaron cuantitativamente los diferentes aspectos de los ítems mediante una escala Likert de 10 categorías. De este modo, se solicitó a los jueces expertos que valorasen los diferentes aspectos relativos a: a) representatividad, o grado en que el ítem tal y como está planteado es el mejor de los posibles; y b) calidad técnica o grado en que el ítem, tal y como está redactado, no induce a error por sesgos gramaticales.

En primer lugar, se realizó la media de las valoraciones cuantitativas con respecto a la representatividad y calidad técnica de los ítems, sin adoptarse ninguna medida adicional, al obtener todos los ítems puntuaciones medias superiores a seis (Osterlind, 1989).

Con respecto a las valoraciones cualitativas, todos los jueces expertos coinciden afirmando que el instrumento CACEF cumple con el grado de comprensión, adecuación y definición requeridas para su uso, aunque consideran necesario un exhaustivo entrenamiento por parte de los observadores que les permita registrar las mismas conductas al analizar los videos, debido a la subjetividad de las conductas.

Dicha idea es apoyada por otro juez experto, quien indica que para registrar las conductas de agresión verbal y gestual, se ha de conocer el contexto en el que se producen, debido a que existen ciertas expresiones entre los alumnos del grupo que pueden ser interpretadas de forma incorrecta por parte de los observadores. Por otro lado, en la Tabla 3 se puede observar la valoración del instrumento de observación se completa con la prueba V de Aiken (Merino y Livia, 2009), donde se aprecia que los resultados obtenidos para cada uno de los ítems tanto en representatividad como calidad técnica, son muy elevadas, superiores a los mínimos propuestos por Penfield y Giacobbi (2004), tanto en la escala global como en cada una de las conductas.

Tabla 3. Valoración cuantitativa de los expertos mediante la prueba V de Aiken

|  |  |  |
| --- | --- | --- |
|  | V de Aiken | |
| Ítems – Conductas | Representatividad | Calidad técnica |
| 1. Agredir verbalmente | 0,92 | 0,77 |
| 2. Agredir físicamente | 0,97 | 0,94 |
| 3. Agredir gestualmente | 0,91 | 0,77 |
| 4. Interrumpir o molestar a los demás | 0,94 | 0,91 |
| 5. Maltratar el material o las instalaciones | 0,88 | 0,88 |
| 6. Hacer trampas, engañar en los juegos y tareas | 0,88 | 0,86 |
| 7. Excluir a compañeros | 0,91 | 0,77 |
| 8. Quejarse | 0,91 | 0,88 |
| 9. No seguir las indicaciones del profesor | 0,88 | 0,80 |
| Global | 0,91 | 0,84 |

Fuente: elaboración propia.

*Fiabilidad*

La fiabilidad es definida como reproducibilidad de una medida (Thomas, Silverman, y Nelson, 2015). En este estudio se calculó la fiabilidad interobservadores y la fiabilidad intraobservador mediante la prueba estadística Coeficiente de Correlación Intraclase (CCI).

- Evaluación de la fiabilidad interobservadores

Para evaluar la concordancia entre observadores, se contó con la colaboración de tres nuevos observadores formados en la observación de conductas que alteraban el clima de aula en Educación Física. En este proceso, se explicó detenidamente el instrumento y se analizaron varias sesiones, tras las cuales se hacía una puesta en común de los acuerdos en el número y tipos de conductas observadas. En aquellas conductas en las que no existía acuerdo, se volvía a analizar la escena y se exponían las causas por las que unos observadores sí habían detectado alguna conducta y por qué otros no, elaborando un protocolo para sucesivas observaciones. Tras un número aproximado de cinco sesiones de entrenamiento se procedió a analizar el grado de acuerdo entre los observadores, para lo cual se empleó la prueba estadística denominada Coeficiente de Correlación Intraclase, obteniendo un valor global de 0,89, considerado como muy aceptable.

- Evaluación de la fiabilidad intraobservador

Por otra parte, se estudió la concordancia intraobservador, a través de la prueba test-retest, para la cual se utilizaron dos sesiones de Educación Física de 45 minutos de duración cada una de ellas, una con alumnos de sexto de Educación Primaria y otra con alumnos de tercero de la ESO. De este modo, tres nuevos investigadores previamente formados en la valoración de las conductas que alteran el clima del aula en las clases de Educación Física visualizaron ambas sesiones en el mismo día, anotando las diferentes conductas que fueron observando y volvieron a repetir este proceso tras un intervalo de una semana. Finalmente, se realizó el cálculo del coeficiente de correlación intraclase, siendo la cifra media de .98 lo que indica un grado de concordancia adecuado.

*Estadísticos descriptivos y consistencia interna*

Las características de los ítems que forman el CACEF fueron analizadas en una muestra de 10 sesiones de Educación Física (Tabla 4), comprobando si el alfa de la escala aumentaba con la eliminación de algún ítem. El coeficiente de correlación corregido ítem-total (CCIT-c) ≥ 0,30 y desviación típica (DT) > 1. Por lo tanto, el ítem siete “excluir a compañeros”, fue eliminado de la hoja de observación, debido a los bajos valores presentados. Como se puede comprobar, y según las recomendaciones de Bollen y Long (1994), los índices de asimetría y curtosis son próximos a 0 y < 2 en la mayoría de los ítems, exceptuando el ítem ocho “quejarse”, que presentó unos valores elevados, y el ítem siete “excluir a compañeros” que fue eliminado. Por lo tanto, la escala final estuvo formada por ocho conductas, mostrando un Alfa de Cronbach de 0,73, valor aceptable para este tipo de investigaciones.

Tabla 4. Estadísticos descriptivos y consistencia interna

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Escala: | M | DT | CCIT-c | α sin ítem | Asimetría | Curtosis |
| 1. Agredir verbalmente | 3,07 | 2,14 | 0,66 | 0,68 | 1,19 | 1,54 |
| 2. Agredir físicamente | 3,14 | 2,27 | 0,50 | 0,69 | 1,21 | 2,13 |
| 3. Agredir gestualmente | 0,79 | 0,99 | 0,51 | 0,71 | 1,68 | 2,25 |
| 4. Interrumpir o molestar a los demás | 14,25 | 8,09 | 0,75 | 0,62 | 0,08 | -0,97 |
| 5. Maltratar el material o las instalaciones | 0,68 | 0,90 | 0,49 | 0,71 | 1,35 | 1,29 |
| 6. Hacer trampas, engañar en los juegos | 13,14 | 4,58 | 0,41 | 0,69 | 0,29 | -0,51 |
| 7. Excluir a compañeros | 0,36 | 0,73 | 0,05 | 0,73 | 2,37 | 2,91 |
| 8. Quejarse | 1,93 | 1,36 | 0,34 | 0,71 | 0,58 | -0,635 |
| 9. No seguir las indicaciones del profesor | 14,21 | 7,78 | 0,72 | 0,62 | 0,15 | -0,303 |

Fuente: Elaboración propia.

*Análisis Factorial Exploratorio*

Con objeto de agrupar las ocho conductas restantes en diferentes categorías o dimensiones se ha realizado un análisis factorial exploratorio con máxima verosimilitud, teniendo en cuenta lo expuesto anteriormente en este trabajo en relación a la consistencia interna (Tabla 4), eliminando el ítem siete “excluir a los compañeros”. Entre las salidas obtenidas se encuentra el índice KMO (Kaiser-Meyer-Olkin) que compara los coeficientes de correlación lineal con los coeficientes de correlación parcial. Si los coeficientes de correlación parcial entre las variables son muy pequeños, quiere decir que la relación entre cada par de las mismas se debe o puede ser explicada por el resto y por tanto se puede llevar a cabo un análisis factorial.

Calculada la medida KMO, el resultado obtenido de 0,65, dentro del límite a partir del cual el análisis factorial a realizar es adecuado y las conclusiones consistentes. Al desarrollar la prueba de esfericidad de Bartlett, se obtiene un p-valor de 0,000, lo que hace rechazar la hipótesis de que la matriz de correlaciones es una matriz de identidad, confirmándose una de las hipótesis de partida. El determinante de la matriz de correlaciones ofrece un valor muy bajo y muchas de estas correlaciones son significativas, apoyando la idea de que el análisis factorial es adecuado.

Al realizar el análisis factorial exploratorio se obtienen dos factores o dimensiones que explican el 68,039% de la variabilidad total, lo que supone un porcentaje bastante aceptable, obteniendo los grupos de ítems que saturan cada factor una vez rotados. Así se obtiene un primer factor formado por los ítems “Agredir gestualmente”, “Agredir verbalmente”, “Agredir físicamente” y “maltratar el material o las instalaciones” que hacen referencia a la categoría de conductas violentas. El factor dos “no seguir las indicaciones del profesor”, “quejarse”, “hacer trampas, engañar en los juegos y deportes” e “interrumpir o molestar a los demás” hacen referencia a la categoría de indisciplina o antisociales. De este modo, las conductas se agrupan en dos factores o categorías: conductas violentas y conductas de indisciplina o desinterés (Tabla 5), siguiendo la línea de los autores Kulinna y col. (2003).

Tabla 5. Factores obtenidos y saturación de los ítems en el factor

|  |  |  |  |
| --- | --- | --- | --- |
| Conductas Violentas | Saturación | Conductas de indisciplina o desinterés | Saturación |
| Agredir verbalmente | 0,52 | Interrumpir o molestar a los demás | 0,94 |
| Agredir físicamente | 0,59 | Hacer trampas, engañar en los juegos | 0,37 |
| Agredir gestualmente | 0,99 | Quejarse | 0,27 |
| Maltratar el material o las instalaciones | 0,50 | No seguir las indicaciones del profesor | 0,60 |

Fuente: elaboración propia.

*Análisis Factorial Confirmatorio*

Con el objetivo de confirmar las dimensiones anteriores, se aplicaron modelos de ecuaciones estructurales, evaluando la estructura factorial del instrumento mediante AFC utilizando el programa Amos 18.0, obteniendo los parámetros a través del método de máxima verosimilitud. Para evaluar el ajuste de los modelos confirmatorios se han utilizado varios índices, absolutos y relativos, como recomiendan Bentler (2007) y Markland (2007). De este modo, para los índices absolutos se utilizó el valor *p*, asociado con el estadístico chi cuadrado (*X2*), así como la ratio entre el índice de *X2* y los grados de libertad (gl), considerando los valores por debajo de 2,0 como indicadores de muy buen ajuste de modelo (Tabachnik y Fidell, 2007). Asimismo, se ha calculado el GFI (índice de bondad de ajuste), y entre los índices relativos el NFI (*índice de ajuste normalizado*), el NNFI (*índice de ajuste no normativo*) y el CFI (*índice de ajuste comparativo*), cuyos valores deben ser iguales o superiores a 0,90 para considerar aceptable el ajuste de un modelo (Kaplan, 2000). Autores como Kline (2005) recomiendan la utilización del RMSEA (error de aproximación cuadrático medio) y, según Steiger (2007), un valor menor a 0,07 indicaría un buen ajuste.

A continuación, en la Figura 2, se presentan los resultados del AFC correspondiente al modelo de ocho conductas que alteran la convivencia en las clases de Educación Física. Los índices de bondad mostraron un ajuste satisfactorio: *X2* = 26,78, gl = 19, *X2*/gl = 1,40, GFI = 0,84, CFI = 0,88, RMSEA = 0,12.

Agredir físicamente

Agredir verbalmente

Agredir gestualmente

Maltratar el material o las instalaciones

Hacer trampas, engañar en juegos y deportes

Interrumpir o molestar a los demás

No seguir las indicaciones del profesor

Quejarse

.74

.62

.73

.59

.44

.83

.87

.34

.56

.39

.54

.35

.19

.70

.75

.12

.80

Figura 2. Diagrama de la trayectoria (Path Diagram) del AFC, con pesos estandarizados y errores de medición de cada una de las conductas del CACEF

**Discusión**

El objetivo del presente estudio ha sido diseñar y validar un instrumento de observación para el análisis de las conductas que alteran la convivencia en las clases de Educación Física. Los resultados obtenidos en la investigación muestran como el diseño y elaboración de una hoja de observación requiere de diversas fases, así como colaboración de jueces expertos (Carretero-Dios y Pérez, 2005; Cruz y col., 2006; Kulinna y col., 2003).

El número de jueces expertos consultado fue superior al empleado en otros trabajos similares (Baena, Granero, Brancho, y Pérez, 2012), realizando importantes aportaciones para la mejora de la hoja de observación, girando en torno a la definición, agrupación y objetividad de las conductas observadas. Dichas contribuciones de tipo cualitativo son indispensables en el desarrollo de un instrumento (Carretero-Dios y Pérez, 2005; Ortega, Jiménez, Palao, y Sainz de Barranda, 2008). Por otro lado, los valores cuantitativos otorgados por el panel de expertos fueron muy elevados, y se reflejan en la prueba V de Aiken.

En este proceso de validación del instrumento, se comprobó la fiabilidad mediante la consistencia interna, medida a través del coeficiente alfa de Cronbach, obteniendo un resultado para el total de la hoja de observación de 0,73, valor aceptable para este tipo de investigaciones (Morales Vallejo, 2007), y similar a la obtenida por otros autores en este tipo de instrumentos (Baena y col., 2012; Cruz y col., 2006; Díaz y Esteban, 2010). Del mismo modo se realizó la prueba estadística Coeficiente de Correlación Intraclase, para analizar la fiabilidad interobservadores, obteniendo un valor de 0,89; y la fiabilidad intra-observador, obteniendo un resultado de .98; ambos valores considerados como muy aceptables, y en concordancia a los obtenidos en otras investigaciones similares (Cruz y col., 2006; Gálvez-Nieto, Vera-Bachman, Cerda, y Díaz, 2016).

Por último, los análisis psicométricos y su correspondiente análisis factorial confirmatorio basado en ecuaciones estructurales apoyan la validez factorial y la fiabilidad del instrumento, formado por ocho conductas que alteran la convivencia en la clases de Educación Física, agrupadas en dos dimensiones (conductas violentas y conductas de indisciplina o desinterés). De este modo, el número conductas a evaluar y de ítems que contiene este instrumento se halla en un término medio con respecto al resto de instrumentos desarrollados hasta el momento. Contiene un número inferior a los instrumentos creados por Cruz y col. (2006), Esteban y Díaz (2010), Auné, Abal, y Attorresi (2016), formado por 59 conductas, pero mayor en comparación con los empleados en los estudios de Escartí y col. (2005) y de Gras y Martínez (2002).

De este modo, el instrumento de observación de conductas que alteran la convivencia en las clases de Educación Física ha obtenido altos valores en los índices de fiabilidad (inter e intraobservador y consistencia interna) y validez, para aquellas clases de Educación Física filmadas en vídeo, así como en el análisis factorial exploratorio y confirmatorio. Por lo tanto, este instrumento permitirá analizar de forma cuantitativa, a través de la observación indirecta, el número y tipología de conductas que se producen en las aulas, que permitirá al docente, conocer el nivel de convivencia en sus clases y comparar de manera objetiva con las aulas de otros centros.

Se puede concluir, que tras este procedimiento encaminado a la creación y validación de un instrumento para registrar las conductas que alteran la convivencia en las clases de Educación Física, los resultados son satisfactorios. Se trata de un sistema de categorías formado por ocho ítems agrupados en dos dimensiones, considerándose exhaustivas y mutuamente excluyentes, y que permite recoger un amplio abanico de conductas que alteran la convivencia, abarcando desde una visión integradora, la tipología de comportamientos que hacen que el ambiente de clase se vea perjudicado.

**Referencias**

Anguera, M. T. (2010). Posibilidades y relevancia de la observación sistemática por el profesional de la psicología. *Papeles del Psicólogo, 31*(1), 122-130.

Anguera, M. T.; Blanco, A.; Hernández, A., y Losada, J. L. (2011). Diseños observacionales: ajuste y aplicación en psicología del deporte. *Cuadernos de Psicología del Deporte, 11*(2), 63-76.

Auné, S. E.; Abal, F. J., y Attorresi, H. F. (2016). Diseño y construcción de una Escala de Conducta Prosocial para adultos. *Revista Iberoamericana de Diagnóstico y Evaluación Psicológica, 42*(2), 15-25.

Baena, A.; Granero, A.; Brancho, C., y Pérez, F. J. (2012). Versión española del Sport Satisfaction Instrument (SSI) adaptado a la Educación Física. *Revista de Psicodidáctica, 17*(2), 377-396.

Barnes, T. N.; Smith, S. W., & Miller, M. D. (2014). School-based cognitive-behavioral interventions in the treatment of aggression in the United States: A meta-analysis. *Aggression and Violent Behavior, 19*, 311-321.

Bentler, P. M. (2007). On tests and indices for evaluating structural models. *Personality and Individual Differences, 42*, 825-829.

Bollen, K. A., & Long, J. (1994). *Testing structural equation models*. Newbury Park, CA: Sage.

Borràs, A.; Palou, P., & Ponseti, F. X. (2001). Evaluación del fairplay y conductas violentas durante la práctica deportiva de estudiantes de magisterio en Educación Física. En J. F. Campos, S. Llanas y R. Aranda (Coords.), *Nuevas aportaciones al estudio de la actividad física y el deporte.* II Congreso de CAFyD. Valencia: Marí Montañana.

Carretero-Dios, H., y Pérez, C. (2005). Normas para el desarrollo y revisión de estudios instrumentales. *International Journal of Clinical and Health Psychology, 5*(3), 521-551.

Cruz, J.; Capdevila, L.; Boixados, M.; Pintanel, M.; Alonso, C.; Mimbrero, J., y Torregrosa, M. (1996). Identificación de conductas, actitudes y valores relacionados con el fairplay en deportistas jóvenes. En *Valores sociales y deporte. Fair Play versus violencia. Investigaciones en ciencias del deporte* (pp. 37-67). Madrid: Consejo Superior de Deportes.

Díaz, A., & Esteban, R. (2010). Physical Education Classroom Management Instrument. Adaptación transcultural. *Cuadernos de Psicología del Deporte, 10*, 73-78.

Escartí, A.; Pascual, C., y Gutiérrez, M. (2005). *Responsabilidad personal y social a través de la educación física y el deporte*. Barcelona: Graó.

Fernández del Río, A. B., & Barreiras Arias, A. J. (2015). Stop-Motion: Estrategia de Imagen y Comunicación para la mejora de la convivencia. *Revista de Estudios e Investigación en Psicología y Educación, 13*, 191-195.

Gálvez-Nieto, D.; Vera-Bachman, D.; Cerda, D., y Díaz, R. (2016). Escala de victimización entre adolescentes a través del teléfono móvil y de internet: Estudio de validación de una versión abreviada en estudiantes chilenos. *Revista Iberoamericana de Diagnóstico y Evaluación Psicológica, 41*(1), 16-27.

García, L. (2010). Los planes de convivencia de centro. Reflexiones y orientaciones. En J. J. Gázquez, y M. C. Pérez (Eds.), *La convivencia escolar. Aspectos psicológicos y educativos* (pp. 73-79). Granada: GEU.

Gras, M., y Martínez, G. (2002). Las primeras manifestaciones de la conducta antisocial en la escuela. *Revista Electrónica Interuniversitaria de Formación del Profesorado, 5*(4), 15.

Goyette, R.; Dore, R., & Dion, E. (2000). Pupils’ misbehaviour and the reactions and causal attributions of physical education student teachers: A sequential analysis. *Journal of Teaching in Physical Education, 20*, 3-14.

Hemphill, M. A.; Templin, T. J., & Wright, P. M. (2015). Implementation and outcomes of a responsibility-based continuing professional development protocol in physical education. *Sport, Education and Society*, *20*(3), 398-419.

Hernández, A.; Díaz, F., y Morales, V. (2010). Construcción de una herramienta observacional para evaluar las conductas prosociales en las clases de Educación Física. *Revista de Psicología del Deporte, 19*(2), 305-318.

Kaplan, D. (2000). *Structural equation modeling: Foundations and extensions*. Thousand Oaks, CA: Sage.

Kuhn, T. M.; Ebert, J. S.; Gracey, K. A.; Chapman, G. L., & Epstein, R. A. (2015). Evidence-based interventions for adolescents with disruptive behaviors in school-based settings. *Child and Adolescent Psychiatric Clinics of North America, 24*(2), 305-317.

Kulinna, P. H.; Cothran, D., & Recualos, R. (2003). Development of an instrument to measure disruptive behaviour. *Measurerement in Physical Education and Exercice Science 7*(1), 25-41.

Kline, R. B. (2005). *Principles and practice of structural equation modeling* (2nd Edition ed.). New York: The Guilford Press.

Landeta, J. (1999). *El método Delphi. Una técnica de previsión para la incertidumbre.* Ariel. Barcelona.

Markland, D. (2007). The golden rule is that there are not golden rules: A commentary on Paul Barrett’s recommendations for reporting model fit in structural equation modelling. *Personality and Individual Differences, 42*, 851-858.

Merino, C., y Livia, S. (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*(1), 169-171.

Morales Vallejo, P. (2007). *El contraste de medias. Diferencia entre medias de muestras relacionadas*. Madrid: Universidad Pontificia Comillas.

Ortega, E.; Jiménez, J. M.; Palao, M., y Sainz de Barranda, P. (2008). Diseño y validación de un cuestionario para valorar las preferencias y satisfacciones en jóvenes jugadores de baloncesto. *Cuadernos de Psicología del Deporte, 8*(2), 39-58.

Osterlind, S. J. (1989). *Constructing test items*. Boston: Kluwer.

Palou, P.; Borràs, P. A.; Ponseti, F. X.; Vidal, J., y Torregrosa, M. (2007). Intervención para la promoción de actitudes de fairplay en futbolistas cadetes. *Apunts, 89*, 15-22.

Penfield, R. D., & Giacobbi, P. R. (2004). Applying a score confidence interval to Aiken's item content-relevance index. *Measurement in Physical Education and Exercise Science, 8*([4](http://www.tandfonline.com/toc/hmpe20/8/4)), 213-225.

Ruiz, L. M.; Rodríguez, P.; Martinek, T.; Schilling, T.; Durán, L. J., y Jiménez, P. (2006). El proyecto esfuerzo: un modelo para el desarrollo de la responsabilidad personal y social a través del deporte. *Revista de Educación, 341*, 933-958.

Steiger, J. H. (2007). Understanding the limitations of global fit assessment in structural equation modelling. *Personality and Individual Differences, 42*, 893-898.

Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics* (5th ed.). New York: Allyn and Bacon.

Thomas, J. R.; Silverman, S., & Nelson, J. (2015). *Research methods in physical activity*, 7E. Campaign IL: Human kinetics.

Urbina, C.; Simón, C., y Echeita, G. (2011). Concepciones de los profesores acerca de las conductas disruptivas: análisis a partir de un marco inclusivo. *Infancia y Aprendizaje, 34*(2), 205-217.

Whitley, M. A.; Forneris, T., & Barker, B. (2015).The reality of sustaining community-based sport and physical activity programs to enhance the development of underserved youth: Challenges and potential strategies. *Quest, 67*(4), 409-423.