

Percepción del alumnado universitario de Educación Física sobre la calificación dialogada

Perception of university students of Physical Education on the dialogue mark

*Fernando Manuel Otero Saborido, **Francisco José Pozuelo-Estrada, ***Constanza Palomino-Devia
*Universidad Pablo de Olavide (España), **Universidad de Huelva (España), ***Universidad de Tolima (Colombia)

Resumen. La calificación dialogada (CD) es una estrategia de evaluación formativa y compartida de la que existe poca evidencia científica. Por ello, el objetivo principal de este estudio era analizar la percepción del alumnado universitario sobre el uso de la calificación dialogada dentro del modelo flipped learning. Para ello, 144 estudiantes del Grado de Ciencias de la Actividad Física y del Deporte realizaron un proyecto de trabajo sobre modelos de iniciación deportiva. Tras un proceso de feedback y evaluación continua sobre la tarea a lo largo de tres semanas, el alumnado compartió una sesión de CD con su docente para la calificación del trabajo. Posteriormente, el alumnado fue preguntado a través de un cuestionario con preguntas semiestructuradas por la sesión de CD. El análisis de las respuestas con el software con Atlas.ti reflejó la percepción positiva del alumnado hacia la CD. Los estudiantes consideran que beneficia al aprendizaje y mejora la interacción con el docente. Así mismo, opinan que la CD habría empeorado si en lugar de utilizar rasgos cualitativos se hubiera negociado sobre calificaciones numéricas.

Palabras clave: Evaluación formativa, retroalimentación, aprendizaje autorregulado, flipped learning.

Abstract. The dialogue score (DM) is a formative and shared assessment strategy for which there is little scientific evidence. Therefore, the main objective of this study was to analyze the perception of university students on the use of the DM within the flipped learning model. To do this, 144 students from the Degree in Physical Activity and Sports Sciences carried out a work project on sports initiation models. After a process of feedback and continuous evaluation of the task over three weeks, the students shared a DS session with their teacher to grade the work. The analysis of the interviews with the Atlas.ti software reflected the positive perception of the students towards DM. Students consider that it benefits learning and improves interaction with the teacher. Likewise, they believe that DM would have worsened if instead of using qualitative traits it had been negotiated on numerical ratings

Keywords: Formative assessment, feedback, self-regulated learning, flipped learning.

Introducción

Aunque en la actualidad cada vez es mayor el uso formativo y compartido de la evaluación (Barrientos Hernán et al., 2018; Bores-García, Hortigüela-Alcalá, Hernando-Garijo, et al., 2020; López Pastor et al., 2015), aún es preponderante en las prácticas educativas la tradición de la evaluación entendida como un proceso exclusivamente técnico en detrimento de su carácter social, ético y educativo (Santos Guerra, 2003). Este predicamento de la evaluación asumida únicamente como calificación sucede a pesar de la contribución para el aprendizaje de una evaluación formativa y compartida. En este sentido, se entiende que una evaluación formativa y compartida tiene como finalidad principal «mejorar los procesos de enseñanza y

aprendizaje» (López Pastor & Pérez Pueyo, 2017). En ese carácter de contribución al aprendizaje es determinante la participación del alumnado en el proceso. Relacionado con esta interacción con el discente, el feedback es uno de los procesos que mayor fiabilidad científica tenemos en lo que se refiere a su contribución sobre el aprendizaje (Brooks et al., 2021; Brooks, Carroll, et al., 2019; Brooks, Huang, et al., 2019; Hattie & Clarke, 2018; Wisniewski et al., 2020). En la etapa y la materia que nos ocupa en el presente estudio, la literatura especializada ya ha evidenciado tanto la ganancia de aprendizaje que genera la implicación del alumnado de Educación Física (EF) de la enseñanza universitaria en su proceso de evaluación (Eather et al., 2017; Macken et al., 2020; Soytürk, 2019), como la percepción positiva que tiene de participar en estos procesos (López-Pastor et al., 2016; Martos-García et al., 2017; Romero Martín et al., 2015).

López-Pastor, González-Pascual y Barba-Martín (2005) ya arrojaron luz con una aclaración terminológica entre los diferentes grados de participación del alumnado

en la evaluación: autoevaluación, coevaluación, evaluación compartida, autocalificación y la calificación dialogada que será objeto de estudio en el presente trabajo. En el caso de la autoevaluación, entendida proceso de reflexión individual, la evidencia científica señala una asociación muy determinante entre los procesos vinculadas a la autoevaluación (autorregulación, autoeficiencia, motivación...) y los resultados de aprendizaje en el alumnado de la enseñanza universitaria (Bartimote-Aufflick et al., 2016). Si nos centramos en el alumnado de EF como el del presente estudio, diferentes estudios señalan que la percepción del uso de la autoevaluación en el proceso es muy positiva en lo que se refiere a la aportación al proceso de enseñanza aprendizaje en aspectos metodológicos como la organización del grupo-clase, el uso de los espacios o la transparencia en la evaluación en el caso de la enseñanza universitaria (Asún-Dieste et al., 2020; López-Pastor et al., 2016; Macken et al., 2020).

En el caso de los procesos de 'coevaluación', la revisión sistemática de Bores-García, Hortigüela-Alcalá, González-Calvo y Barba-Martín (2020) sobre este mecanismo de participación del alumno en EF selecciona siete trabajos realizados en la enseñanza universitaria objeto de la presente investigación. Con una mezcla de enfoques cualitativos y cuantitativos, se obtienen resultados significativos que podrían resumirse en un aumento de la autoeficiencia del alumnado, la motivación y del desarrollo de las destrezas docentes cuando se utilizan mecanismos de coevaluación.

Los estudios sobre 'autocalificación', entendida como el proceso a través del cual cada alumno fija la calificación que cree merecer (López Pastor & Pérez Pueyo, 2017), son muy escasos. No obstante, en el caso de las enseñanzas no universitarias parece que los estudiantes que se autocalifican obtienen mejores resultados en las evaluaciones realizadas por los docentes que aquellos estudiantes que no se autocalificaron previamente (Sanchez et al., 2017). En el caso de la enseñanza universitaria en EF, los estudios hallados demuestran que existe una alta correlación entre autocalificación del alumnado, la calificación con el docente y la calificación negociada entre estudiantes y docentes (López-Pastor et al., 2012)

Por otra parte, la 'evaluación compartida' es entendida como un proceso de reflexión y diálogo entre docentes y alumnado sobre su proceso de aprendizaje sin que exista calificación. Existen ejemplos en la enseñanza universitaria de un uso sistematizado y cíclico de la evaluación compartida con el caso del modelo

flipped learning (Otero-Saborido, 2020). En el caso de la 'calificación dialogada', objeto de la presente investigación, puede considerarse como un paso más que la evaluación compartida ya que, en este caso, a diferencia de la evaluación compartida, sí existe calificación. La evidencia científica sobre estos procesos de negociación de la calificación en la enseñanza universitaria es escasa (López-Pastor et al., 2012). Desde un punto de vista terminológico, el término ha sido conceptualizado como «proceso por el cual alumnado y profesorado dialogan sobre la calificación definitiva y la acuerdan, en base a los criterios de calificación previamente establecidos» (López Pastor & Pérez Pueyo, 2017). En estos términos, Gómez-Ruiz y Quesada (2020) acuñan el concepto de 'calificación compartida' con igual significado que el de calificación dialogada. Sobre este aspecto McArthur (2019) ya plantea sus reservas sobre la calificación numérica en la enseñanza universitaria y propone liberar la evaluación de esos «esquemas rígidos» para construir una evaluación al servicio de la justicia social. Por ello, es de notable interés científico el estudio de inclusión/exclusión de la calificación numérica dentro de estos procesos de negociación (calificación dialogada/calificación colaborativa). Diferentes autores han señalado el perjuicio de la inclusión de la calificación numérica en estos procesos de diálogos y negociación y de la evaluación en general (Butler, 1988; Calderón & MacPhail, 2018; Ruiz Martín, 2020). Dado este planteamiento teórico, el objetivo principal de este estudio era analizar la percepción del alumnado universitario sobre el uso de la calificación dialogada dentro del modelo *flipped learning*. Así mismo, de forma más concreta dentro de este fin, parecía necesario conocer la opinión del alumnado sobre la inclusión/exclusión de la calificación numérica dentro de estos procesos de diálogo.

Metodología

Participantes

Se partió de una muestra inicial de 144 estudiantes universitarios del Grado de Ciencias de la Actividad Física y del Deporte de la Universidad Pablo de Olavide de Sevilla. Como criterio de inclusión se consideró estar matriculado en la asignatura Fundamentos de la Educación Física y del Deporte de primero del Grado de dicha titulación durante el curso 2020-2021. Como criterios de exclusión se aplicaron dos. En primer lugar, que el estudiante no hubiera cumplimentado algunos de

los cuestionarios del proyecto de trabajo que duró tres semanas. En segundo lugar, que no hubiera participado en la sesión de calificación dialogada. Tras la aplicación de estos criterios la muestra quedó configurada por un total 109 estudiantes (32 alumnas y 72 alumnos) con una edad media 19,7 años. Todos los participantes fueron informados que la recogida de información tenía fines académicos y que se guardaba el anonimato.

Procedimiento

El alumnado debía diseñar para un contexto y modalidad deportiva concreta seis sesiones de un modelo de iniciación deportiva entre los siguientes: tradicional, Educación Deportiva, Comprensivo y Cooperativo. A lo largo de tres semanas, y siguiendo el modelo de flipped learning, el alumnado fue trabajando en la creación de este modelo. Durante la primera semana, tras el trabajo previo con el visionado del video y la lectura de las monografías, se le asignó a cada grupo un caso práctico que comenzaron a diseñar. Previamente, se publicó y explicó la rúbrica diseñada por el equipo docente de la asignatura. El trabajo a valorar se incluía dentro de los contenidos de la guía docente de la asignatura. Los docentes se aseguraron que el alumnado comprendía cada una de las dimensiones y los diferentes niveles de desempeño del instrumento. La rúbrica contenía seis dimensiones y cuatro niveles de desempeño descritos cualitativamente en cada de las dimensiones. Durante la segunda semana, el alumnado recibió el feedback docente sobre el transcurso del trabajo tanto en las sesiones en el aula como en las sesiones en los espacios deportivos. En estas últimas, el alumnado tenía que implementar situaciones reales del modelo deportivo asignado. En este caso, los estudiantes tuvieron la posibilidad de reajustar sus situaciones motrices en base a la retroalimentación docente y de los compañeros. Durante la última semana, cada grupo expuso su trabajo práctico y, posteriormente, en una sesión específica, se realizó una calificación dialogada donde docente y alumnado de cada grupo negociaban los niveles de desempeño de su trabajo en las diferentes dimensiones. Posteriormente, el alumnado completó la preguntas abiertas escrita donde se le preguntaba por la calificación dialogada.

Así mismo, a lo largo de las tres semanas el alumnado realizó una estimación sobre el tiempo de dedicado al proyecto de trabajo (tabla 1). A nivel general, estimó dedicarle un total de 12 horas de trabajo (772,52 minutos). De esa suma unas 6 horas (392,65 minutos) corresponden a las sesiones presenciales y unas 5 horas

(329,87 minutos) al trabajo no presencial. Dentro de la carga no presencial las estimaciones indican un promedio dedicación de casi dos horas (109,96 minutos) en cada una de las 3 semanas que duró el proyecto. Diferenciando el trabajo no presencial por tipología de actividades la preparación del trabajo final es lo que el alumnado estimó dedicarle más tiempo con un promedio de 49,27 minutos semanales seguido de la lectura de textos (22,63 minutos semanales)

Tabla 1.
Tiempo (minutos) promedio de trabajo no presencial y presencial a lo largo de las tres semanas del proyecto para los estudiantes incluidos en la muestra

Volumen		Semana1	Semana 2	Semana 3	Total	
		Promedio	Promedio	Promedio	Promedio	Suma
Presencial	EB	0	68,49	66,90	45,13	135,39
	EPD	85,22	85,22	86,81	85,75	257,25
	TOTAL	85,22	153,71	153,71	130,88	392,65
No presencial	LT	36,73	15,89	15,274	22,63	67,90
	VV	25,75	5,85	6,460	12,69	38,07
	CE	17,42	4,75	7,911	10,02	30,08
	TV	0,98	1,25	1,017	1,083	3,249
	PP	19,88	66,6	61,28	49,27	147,8
	CA	8,79	6,66	7,716	7,724	23,17
	OTR	4,85	7,83	7,0265	6,56	19,70
	TOTAL	114,41	108,77	106,69	109,96	329,87
Volumen total		199,63	262,48	260,40	240,84	772,52

Instrumentos

La recogida de información se realizó a través de un cuestionario con preguntas semiestructuradas que facilitarían la consecución del objetivo ‘Analizar la percepción del alumnado universitario sobre el uso de la calificación dialogada dentro del modelo flipped learning’. Las participantes cumplimentaron las preguntas abiertas por escrito a través de la plataforma de la universidad.

Partiendo del objetivo general del estudio, tres expertos (profesorado universitario en activo familiarizado con las metodologías activas) diseñaron una primera propuesta de preguntas abiertas en base al objetivo del estudio. En la primera circulación los expertos pusieron en común sus propuestas individuales. Tras evaluar el grado de contribución de las propuestas y la factibilidad de las mismas, los expertos consideraron que el excesivo número inicial de preguntas para una práctica tan concreta como la calificación dialogada sesgaba el objetivo del estudio. En la segunda circulación se redujeron el número de preguntas consensuándose un modelo final de preguntas abiertas que preguntaban por aspectos favorables y desfavorables de la calificación dialogada. También se cuestionaba de forma abierta qué hubiera pasado si la calificación dialogada se hubiera realizado teniendo como referencia una calificación numérica en lugar de una rúbrica que desglosaba niveles de desempeño cualitativos muy concretos. Así mismo, la recogida de información incluyó una pregunta cerrada que cuestionaba al alumnado sobre qué lugar ocuparía

la práctica de la calificación dialogada en comparación con las otras prácticas de evaluación compartida realizadas durante el curso (autoevaluación, coevaluación y evaluación compartida).

Por último, con una periodicidad semanal, a través de un cuestionario validado (Otero-Saborido et al., 2018) se recogió información (tabla 1) sobre el tiempo de trabajo no presencial dedicado a cada una de las tipologías de actividades a lo largo de todo el proyecto de trabajo.

Sistema de categorías

Al tratarse de un estudio cualitativo de caso que tiene por objeto la comprensión de una situación concreta, se planificó un sistema de categorías previo (González et al., 2020). A modo de pilotaje, 10 respuestas fueron codificadas por los 3 expertos a partir del sistema de categorías previo. Tras esta codificación piloto, el sistema de categorías quedó configurado con 4 grandes dimensiones y diferentes categorías que la integraban (tabla 2).

Tabla 2.
Sistema de dimensiones, categorías y descripción de las mismas en el análisis de contenido de las respuestas

Dimensiones	Categorías	Descripción de la categoría
1. Aspectos favorables	1.1 Metas de aprendizaje	La respuesta incluye aspectos favorables de la calificación dialogada relacionados con el proceso de adquisición de destrezas, habilidades y /o contenidos relacionados
	1.2 Metas de rendimiento	La respuesta incluye aspectos favorables de la calificación dialogada relacionados con el resultado final o la calificación.
	1.3 Interacción docente	La respuesta incluye aspectos favorables de la calificación dialogada relacionados con la relación docente-alumnado
	1.4 Interacción alumnado	La respuesta incluye aspectos favorables de la calificación dialogada relacionados con la relación entre alumnado
2. Aspectos desfavorables	2.1 No (argumentado)	El participante niega la existencia de aspectos negativos en la calificación dialogada argumentado sus respuestas.
	2.2 No (sin argumentar)	El participante niega la existencia de aspectos negativos en la calificación dialogada sin argumentar sus respuestas.
	2.3 Sí (argumentado)	El participante reconoce la existencia de aspectos negativos en la calificación dialogada argumentado su respuesta.
	2.4 Sí (sin argumentar)	El participante reconoce la existencia de aspectos negativos en la calificación dialogada sin argumentar su respuesta
3. Emociones	3.1 Positivas	Los sentimientos del participante se vinculan emociones positivas
	2.2 Negativas	Los sentimientos del participante se vinculan emociones negativas
	2.3 Metas de aprendizaje	Los sentimientos del alumnado se vinculan con el proceso de adquisición de destrezas, habilidades y /o contenidos relacionados
	2.4 Metas de rendimientos	Los sentimientos del alumnado se vinculan con el resultado final o la calificación
	2.5 Interacción docente	Los sentimientos del alumnado se vinculan a aspectos relacionados con la relación docente-alumnado
	2.6 Interacción alumnado	Los sentimientos del alumnado se vinculan a aspectos relacionados con la relación entre alumnado
4. Empeoramiento	4.1 Calificación	En la respuesta se vincula la calificación o resultado final
	4.2 Aprendizaje	En la respuesta se citan aspectos relacionados con el proceso de adquisición
	4.3 Interacción Docente	En la respuesta se citan aspectos relacionados con la interacción docente-alumnado.

Materiales y análisis de contenido

Las respuestas escritas fueron recogidas a través de la plataforma oficial Blackboard Collaborate de la universidad Pablo de Olavide de Sevilla tras el consentimiento de los participantes. Para el análisis de contenido de los datos obtenidos en la investigación se utilizó el software Atlas.ti 8. En primer lugar, se

estableció un mapa de códigos por dimensiones y categorías, incluyendo las relaciones entre ambos. Una vez codificado todo el contenido de las respuestas, se aplicaron sobre las categorías diferentes técnicas de análisis como la tabla-código descriptiva y el análisis de co-ocurrencias tanto a los 109 participantes como a grupos formados dentro de los mismos tal y como se ha realizado en investigaciones con metodología análogas (De-Juanas Oliva et al., 2020). La primera técnica consiste en una recopilación cuantitativa de la frecuencia de los códigos. En el caso del análisis de co-ocurrencias, se establecen la frecuencia de asociaciones entre diferentes subcategorías y/o variables.

Resultados

El fin general de la investigación era necesario 'Analizar la percepción del alumnado sobre la calificación dialogada como práctica de evaluación formativa' Para ello, en primer lugar, se le preguntó al alumnado qué lugar ocupó la calificación dialogada entre las otras prácticas de evaluación formativa y compartida realizadas a lo largo del curso (coevaluación, autoevaluación y evaluación compartida). En segundo lugar, con un formato de respuesta abierta, se les cuestionó sobre aspectos positivos y negativos y cómo se sintieron en la sesión donde negociaron sobre su calificación. Por último, se trasladó al alumnado la pregunta sobre qué hubiera pasado si en vez de negociar su calificación sobre niveles de desempeños cualitativos en cada una de las dimensiones (muy adecuado, adecuado, poco adecuado, nada adecuado) la práctica de calificación dialogada hubiera consistido en negociar calificaciones numéricas.

Sobre la cuestión que preguntaba el lugar que le asignan a la calificación dialogada, la mayoría del alumnado la elige como la mejor práctica de evaluación formativa por delante de la coevaluación, la autoevaluación y la evaluación compartida (gráfico 1). Así lo señala el 64.6 % (73 de los 113) del alumnado. Por el contrario, solo 2 de los 113 estudiantes (el 2,7%) señalaron que es la peor práctica.

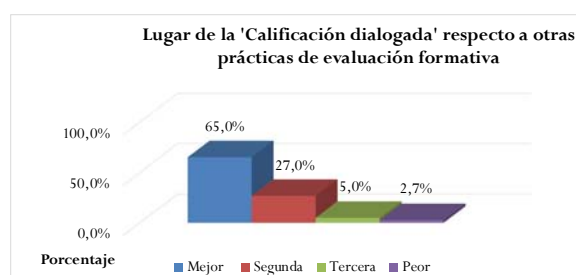


Gráfico 1. Lugar de 'calificación dialogada' respecto a la coevaluación, la evaluación compartida y la autoevaluación

Al mismo tiempo, las respuestas abiertas nos permitieron profundizar por qué la calificación dialogada fue elegida como la mejor práctica. Para ello, se les preguntó al alumnado qué aspectos favorables y desfavorables apreciaron así como los sentimientos que despertó en ello el uso de esta práctica. En el caso de los aspectos favorables las respuestas se agruparon en 4 categorías: a) metas de aprendizaje; b) metas de rendimiento; c) interacción con el docente; d) interacción con el alumnado (gráfico 2).

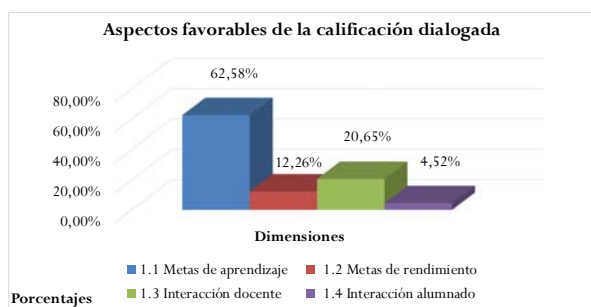


Gráfico 2. Percepción del alumnado sobre los aspectos favorables de la 'calificación dialogada'

Una recopilación de la frecuencia a través de la técnica tabla-código descriptiva del programa Atlas.ti muestra que la mayoría de aspectos favorables de la calificación dialogada se asociaban a la categoría 'metas de aprendizaje'. Es decir, el alumnado entendía que la principal contribución de la calificación dialogada se orientaba a la adquisición de aprendizajes y competencias.

«La calificación dialogada nos ha ayudado a entender qué errores hemos cometido y cómo solucionarlos y ello nos puede ayudar para futuros trabajos en esta asignatura y en otras». (p87)

«Sinceramente pienso que es una manera espléndida de evaluarnos, ya que nos hace reflexionar sobre el trabajo realizado y fomenta la autocrítica del grupo. Analizar el trabajo con los niveles de desempeño de la rúbrica y con nuestro docente, evaluando todos los aspectos del trabajo, lo veo una gran forma tanto de evaluación como de aprendizaje». (p79)

Los aspectos favorables asociados a la 'interacción con el docente' fueron la segunda categoría más mencionada aunque a mucha distancia de las menciones a las metas de aprendizaje.

«Muchos, sobre todo yo resaltaría la interacción entre profesor y alumnos a la hora de destacar cómo ha sido el trabajo realizado» (p79)

Las respuestas que contenían texto relacionado con las 'metas de rendimiento e 'interacción entre alumnado' fueron porcentualmente más bajas. Por último, dentro de los aspectos favorables, cabe resaltar la asociación

obtenida tras la aplicación de la técnica análisis de co-ocurrencias entre las categorías 'metas de aprendizaje' e 'interacción con el docente'. Una gran parte de las respuestas relacionaban que la contribución de la calificación dialogada al aprendizaje contribuía al mismo tiempo a una buena interacción con el docente.

«Hay bastantes aspectos positivos, por ejemplo la cercanía con el docente, es decir nos dejaba responder sus preguntas con un muy buen ambiente entre los alumnos y el docente y los fallos que el veía nos explicaba porque estaban mal... todo esto de muy buena manera y lo más importante es que se aprendía»(p45)

«Veo un punto muy clave el hecho de que el profesor se acerque y ponga justo delante de sus alumnos a «evaluar» la autoevaluación del grupo. Es el momento donde aprendes qué has hecho correctamente para seguir trabajándolo igual o mejor, y además saber qué te ha faltado para mejorar aspectos en un futuro»(p97)

En el caso de los aspectos desfavorables, un 55% de los 110 participantes niega que existan en modo alguno en la calificación dialogada pero sin argumentar su respuesta (gráfico 3). El 18% de los participantes sí argumenta sobre la ausencia de aspectos favorables

«Bajo mi punto de vista no los hay, ya que hubo diálogo entre mis compañeros y el profesor, consiguiendo de este modo un mayor aprendizaje sobre nuestro modelo»(p34).

«Creo que no hay aspectos negativos, ya que es una manera muy justa y beneficiosa de puntuar los trabajos y a veces lo que se echa en falta o nos «mosquea» a los alumnos de las notas es eso, la explicación o aportación por parte del profesor de los fallos o aciertos que hemos tenido, ya que eso ayuda mucho a mejorar o reforzar el grupo de trabajo o al alumno que realiza ese trabajo»(p92)



Gráfico 3. Percepción del alumnado sobre los aspectos desfavorables de la 'calificación dialogada'

Por último, dentro de los aspectos desfavorables un 28% de los encuestados argumentó aspectos desfavorables, aunque la mayoría los plantea de forma hipotética, es decir, consideran que podrían surgir aunque en su experiencia no se manifestaron.

«Pueden existir algunos casos en los que no se llegue

a un acuerdo conforme con el grupo y el profesor, además» (p26)

«Creo que hay gente que podría no ser y quiera poner calificación más alta con intención de sacar mejor nota» (p112)

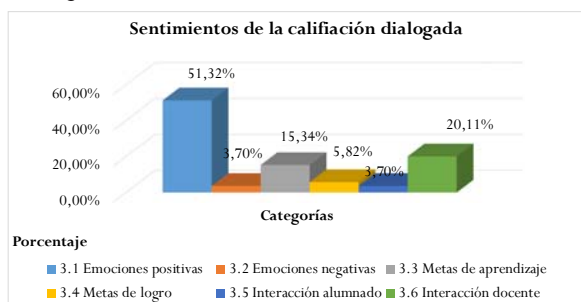


Gráfico 4. Sentimientos del alumnado en la calificación dialogada

Cuando se les preguntó al alumnado cómo se habían sentido, la mayoría de las respuestas (51,32%) del alumnado podrían agruparse bajo sentimientos positivos.

«En general, nos sentimos todo el grupo bastante cómodos y con confianza para dar nuestra opinión o exponer nuestro punto de vista sobre algún tema» (p85).

«Muy tranquilo y cómodo. Todo lo que nos llevemos de ahí es aprendizaje» (p97)

Así mismo, cabe resaltar que gran parte de las emociones positivas que declara el alumnado suele estar vinculadas a la 'interacción con el docente' o a 'metas de aprendizajes' o a ambas categorías al mismo tiempo.

«Me he sentido cómodo gracias a la forma en la que ha planteado el profesor la actividad. Te daba la opción en todo momento de justificar tu nota, y corregirle en aquellas decisiones en las que no estabas de acuerdo, por supuesto, justificándolo.

me sentí bien, ya que solo hubo una cosa que nos corrigió y el grupo lo tomó como crítica constructiva que lo hacía era que aprendieras para futuros trabajos» (p10)

Por último, y dado que en la práctica de calificación dialogada se negoció con el alumnado sobre niveles de desempeño cualitativos y no sobre calificaciones numéricas, se le preguntó al alumnado cómo hubiera transcurrido el diálogo entre docente y alumnado si se hubieran utilizado calificaciones numéricas en lugar de los niveles de desempeño utilizados. Se les dieron tres opciones de respuesta. El diálogo hubiera transcurrido 'igual', habría 'empeorado' o habría 'mejorado' (gráfico 2). El 85% del alumnado (96 de los 114 estudiantes) considera que el diálogo habría 'empeorado' si la negociación se hubiera producido sobre calificaciones numéricas. El 11% considera que el diálogo no habría cambiado si se hubiera optado por números. Sólo 3

estudiantes, el 3,5% del total, considera que el diálogo habría 'mejorado' si se hubieran utilizado calificaciones numéricas.

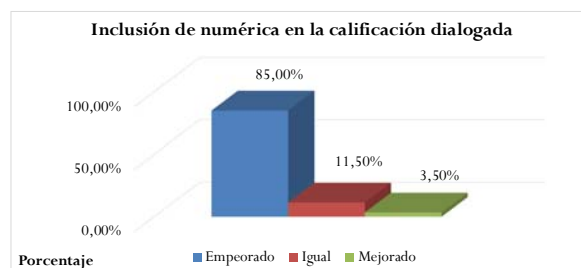


Gráfico 5. Predicción de los estudiantes en el transcurso del diálogo entre docente-alumnado si se hubieran incluido calificaciones numéricas en lugar de niveles de desempeño

El análisis de contenido de las respuestas refleja un gran porcentaje (52,24%) que justifica la inclusión de números como forma de calificar es la principal causa que podría provocar ese empeoramiento (gráfico 6).

«Cuantificar un trabajo siempre «da problemas» a menos que el resultado sea muy positivo, pues hay menos comprensión y si no sabes qué se tiene en cuenta a la hora de evaluar y poner una nota u otra, aparecen las discrepancias y en vez de cercanía y autocorrección, aparece un mínimo rechazo a la asignatura y al profesor» (p1)

«Al utilizar números se hubiera perdido parte de la reflexión pues estos hubieran centrado toda la atención, dejando un poco de lado lo verdaderamente importante que es dialogar, aprender, reflexionar y comprender al 100% la materia» (p18)

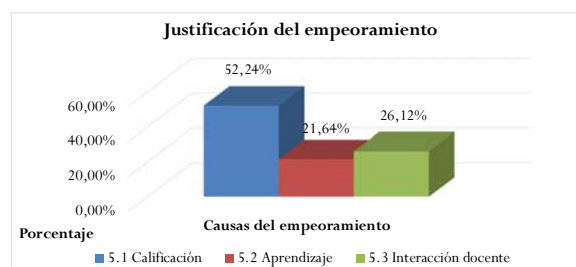


Gráfico 6. Justificación del empeoramiento de la calificación dialogada en el caso de que se hubieran utilizado calificaciones numéricas en lugar de niveles de desempeño

Así mismo, un porcentaje considerable (26,14%) expone que la inclusión numérica habría afectado y/o modificado la interacción con el docente

«El dialogo hubiera sido más a la defensiva por nuestra parte quizás y no escucharíamos tanto la opinión del maestro» (p51)

«Pienso que el diálogo se hubiera empeorado porque nos está poniendo una nota sin darnos ningún motivo» (p70).

Por último, la mención al aprendizaje es igualmente relevante en porcentaje cuando se justifica el empeoramiento del clima en una hipotética inclusión de numérica en la calificación dialogada.

«No justifica la nota por lo que no conoceríamos nuestros fallos y no podríamos mejorar» (p94)

«Hay que dar argumentos y decir las cosas a través de palabras, las razones de porque está bien o porque está mal, que falta y que hay que añadir para mejorar nuestro aprendizaje» (p104)

Por último, se hicieron en análisis grupos estratificados (sexo y vía de acceso a la titulación) sin que se hallaran diferencias en los resultados

Discusión y conclusiones

El fin general de la investigación era necesario 'Analizar la percepción del alumnado sobre la calificación dialogada como práctica de evaluación formativa'. En primer lugar, hay que reseñar que la 'calificación dialogada' como mecanismo de participación del alumnado ha sido escasamente investigada en la literatura especializada y, en menor medida en la enseñanza universitaria (López-Pastor et al., 2012; Sanchez et al., 2017). La principal razón es que estamos ante una estrategia de participación reciente. La conceptualización de esta práctica de evaluación formativa se realizó en primera instancia en 2005 por la Red de Evaluación Formativa y Compartida en Educación (López Pastor et al., 2005). Posteriormente, estos conceptos se han consolidado desde un punto de vista teórico, práctico y empírico a lo largo de diferentes trabajos (Herráz-Sancho, 2017; López Pastor et al., 2007; López Pastor & Pérez Pueyo, 2017; Pérez-Pueyo et al., 2017; Pérez-Pueyo & López-Pastor, 2017). Por tanto, estamos ante una estrategia de participación del alumnado en la evaluación relativamente joven si la comparamos con las primigenias 'autoevaluación' y 'coevaluación'. Así mismo, la 'calificación dialogada' es una consecuencia lógica de los procesos de 'evaluación compartida' o, alternativamente, de los procesos de 'autocalificación' que son, a su vez, relativamente 'jóvenes'. No obstante, si bien aún las investigaciones de calificación dialogada son escasas, sí encontramos experiencias educativas en Primaria, Secundaria y en la enseñanza Universitaria. En el caso de la enseñanza no universitaria, el trabajo de Herráz-Sancho (2017) utilizando la 'calificación dialogada' tras haber realizado previamente con el alumnado 'evaluación compartida' y 'calificación dialogada' o, en el caso de Secundaria, el de Pérez-Pueyo y López-Pastor (2017) utilizando la 'calificación dialogada' para la valoración del trabajo intragrupal con estrategias como la 'Nota Individual Complementada' dentro del Estilo Actitudinal. En el

caso de las enseñanzas universitaria, la experiencia de Pérez-Pueyo, Barba-Martín, López-Pastor y Lorente-Catalán (2017) utilizó la 'calificación dialogada' para el reparto de calificaciones dentro del grupo y en la de Santos-Pastor y Fernández-Muñoz (2017) docente y alumnado negociaron una calificación numérica.

Centrándonos en los resultados del presente estudio, la percepción del alumnado respecto al uso de la 'calificación dialogada' es muy favorable ya que identifica esta práctica con metas relacionadas con la mejora del aprendizaje que se vinculan así mismo con una buena interacción con el docente. En la misma línea, el alumnado destaca las emociones positivas que les provoca la sesión de 'calificación dialogada' realizada con el docente. Si bien la percepción del alumnado es muy positiva en otras prácticas de evaluación formativa como la 'autoevaluación' o la 'coevaluación' (Asún-Dieste et al., 2020), los estudios y/o experiencias revisados sobre calificación dialogada no tuvieron como objetivo analizar la percepción del alumnado respecto a esta práctica (López-Pastor et al., 2012; Pérez-Pueyo et al., 2017; Sanchez et al., 2017; Santos-Pastor & Martínez-Muñoz, 2017). Por ello, no se pueden establecer datos comparativos a este respecto con los datos de este estudio.

Por último, dentro del objetivo general de este estudio se pretendía de forma más concreta 'conocer la opinión del alumnado sobre la inclusión/exclusión de la calificación numérica dentro de estos procesos de diálogo'. El 85% (figura 4) del alumnado consideró que la sesión de calificación dialogada habría empeorado si, en lugar de negociar niveles de desempeño cualitativos con sus respectivas descripciones en una rúbrica, se hubiera negociado sobre calificaciones numéricas. La mayoría del alumnado atribuye ese perjuicio a que las calificaciones numéricas desvían la atención pasando de unas metas de aprendizaje a unas metas de rendimiento. Sobre estos resultados cabe destacar el estudio Butler (1988) que tenía como objetivo comprobar los efectos que tenía la calificación sobre el aprendizaje. Para ello se diseñó un estudio con tres grupos de alumnado y dos pruebas de evaluación. La primera prueba al principio y la segunda al final de la investigación. Tras la realización de la primera medición, el primer grupo recibió feedback sólo mediante las calificaciones; el segundo grupo recibió comentarios cualitativos y calificaciones mientras que el tercer grupo sólo recibió comentarios. Los resultados de la repetición de la segunda prueba mostraron que sólo el grupo que recibió exclusivamente comentarios cualitativos (sin calificación) mostró interés

por la tarea evaluada y, por tanto, mejoró los resultados de su aprendizaje. El grupo que recibió como feedback las calificaciones no mejoró, pero tampoco el que recibió feedback y calificaciones. En este estudio se evidenció que las calificaciones tienen efectos psicológicos que desvían la atención y, por ende, la motivación de la tarea. En ese sentido, el metanálisis realizado por Kluher y Denisi (1996) sobre más de 6000 tamaños del efecto demuestra que el feedback por sí sólo no es beneficioso para el aprendizaje sino que, en ocasiones, dependiendo sus características como ocurría en estudio de Butler (1988), podría no aportar nada o, incluso, ser perjudicial para el proceso de diálogo y, por tanto, para el aprendizaje.

A modo de conclusión, habría que señalar que, desde la perspectiva del alumnado, la ‘calificación dialogada’ es un proceso beneficioso para el aprendizaje. Así mismo, este mecanismo de participación debería sistematizarse en la práctica docente como consecuencia de otras estrategias previas de evaluación formativas como la ‘evaluación compartida’ o la ‘autocalificación’. De forma natural el alumnado demanda la necesidad de comprender sus calificaciones y, en ocasiones, de discutir las (Landrum, 2019). Igualmente, el alumnado opina que la calificación dialogada habría empeorado si en lugar de utilizar rasgos cualitativos se hubiera negociado sobre calificaciones numéricas. La perspectiva del alumnado sobre la calificación dialogada debería de ser explorada en otros contextos educativos y con otras experiencias metodológicas.

Referencias

- Asín-Dieste, S., Romero-Martín, R., Aparicio-Herguedas, J. L., & Fraile-Aranda, A. (2020). Conducta proxiémica en la formación inicial del profesorado en educación física. *Apunts. Educación Física y Deportes*, 141, 41–48.
- Barrientos Hernán, E., López Pastor, V. M., & Pérez-Brunicardi, D. (2018). ¿Por qué hago evaluación formativa y compartida y/o evaluación para el aprendizaje en EF? La influencia de la formación inicial y permanente del profesorado (Why do I do Formative and Share Assessment and/or Assessment For Learning in Physical Education?). *Retos*, 2041(36), 37–43. <https://doi.org/10.47197/retos.v36i36.66478>
- Bartimote-Aufflick, K., Bridgeman, A., Walker, R., Sharma, M., & Smith, L. (2016). The study, evaluation, and improvement of university student self-efficacy. *Studies in Higher Education*, 41(11), 1918–1942. <https://doi.org/10.1080/03075079.2014.999319>
- Bores-García, D., Hortigüela-Alcalá, D., González-Calvo, G., & Barba-Martín, R. (2020). Peer assessment in physical education: A systematic review of the last five years. *Sustainability (Switzerland)*, 12(21), 1–15. <https://doi.org/10.3390/su12219233>
- Bores-García, D., Hortigüela-Alcalá, D., Hernando-Garijo, A., & González-Calvo, G. (2020). Analysis of student motivation towards body expression through the use of formative and share assessment. *Retos*, 2041(40), 198–208. <https://doi.org/10.47197/RETOS.V1140.83025>
- Brooks, C., Burton, R., van der Kleij, F., Carroll, A., Olave, K., & Hattie, J. (2021). From fixing the work to improving the learner: An initial evaluation of a professional learning intervention using a new student-centred feedback model. *Studies in Educational Evaluation*, 68. <https://doi.org/10.1016/j.stueduc.2020.100943>
- Brooks, C., Carroll, A., Gillies, R. M., & Hattie, J. (2019). A matrix of feedback for learning. *Australian Journal of Teacher Education*, 44(4), 14–32. <https://doi.org/10.14221/ajte.2018v44n4.2>
- Brooks, C., Huang, Y., Hattie, J., Carroll, A., & Burton, R. (2019). What Is My Next Step? School Students’ Perceptions of Feedback. *Frontiers in Education*, 4. <https://doi.org/10.3389/educ.2019.00096>
- Butler, R. (1988). Enhancing and undermining intrinsic motivation: The effects of task-involving and ego-involving evaluation of interest and performance. *British Journal of Educational Psychology*, 58(1), 1–14. <https://doi.org/10.1111/j.2044-8279.1988.tb00874.x>
- Calderón, A., & MacPhail, A. (2018). Two sides of the same coin: Student-teachers’ perceptions of a cyclical self-assessment process. *Infancia, Educación y Aprendizaje*, 3(2), 638–642. <https://doi.org/10.22370/ieya.2017.3.2.793>
- De-Juanas Oliva, Á., García-Castilla, F. J., & Ponce de León Elizondo, A. (2020). El tiempo de los jóvenes en dificultad social: utilización, gestión y acciones socioeducativas. *Revista Española de Pedagogía*, 78(277), 477–495. <https://doi.org/10.22550/rep78-3-2020-05>
- Eather, N., Riley, N., Miller, D., & Jones, B. (2017). Evaluating the effectiveness of using peer-dialogue assessment (PDA) for improving pre-service teachers’ perceived confidence and competence to teach physical education. *Australian Journal of Teacher Education*, 42(1), 69–83. <https://doi.org/10.14221/ajte.2017v42n1.5>
- Gómez Ruiz, M. Á., & Quesada Serra, V. (2020). Análisis de las calificaciones compartidas en la modalidad participativa de la evaluación colaborativa entre docente y estudiantes. *RELIEVE - Revista Electrónica de Investigación y Evaluación Educativa*, 26(1), 1–19. <https://doi.org/10.7203/relieve.26.1.16567>
- González, D. H., López-Pastor, V. M., & Arribas Manrique, J. C. (2020). La Evaluación Formativa y Compartida en contextos de Aprendizaje Cooperativo en Educación Física en Primaria. *Cultura, Ciencia y Deporte*, 15(44), 213–222. <https://doi.org/10.12800/ccd.v15i44.1463>
- Hattie, J., & Clarke, S. (2018). Visible learning: Feedback. In *Visible Learning: Feedback*. <https://doi.org/10.4324/9780429485480>
- Herráz-Sancho, M. (2017). La participación del alumnado en los

- procesos de evaluación en educación Primaria: autoevaluación y evaluación compartida. In V. López-Pastor & A. Pérez-Pueyo (Eds.), *Evaluación formativa y compartida en educación: experiencias de éxito en todas las etapas educativas* (pp. 180–193). Universidad de León. <https://buleria.unileon.es/handle/10612/5999>
- Kluger, A. N., & DeNisi, A. (1996). The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin*, 119(2), 254–284. <https://doi.org/10.1037/0033-2909.119.2.254>
- Landrum, B. (2019). 'See Me as I see Myself': A Phenomenological Analysis of Grade Bump Requests. *Qualitative Research in Education*, 8(3), 315–340. <https://doi.org/10.17583/qre.2019.4329>
- López-Pastor, V. M., Fernández-Balboa, J., Santos Pastor, M. L., & Fraile-Aranda, A. (2012). Students' self-grading, professor's grading and negotiated final grading at three university programmes: analysis of reliability and gradedifference ranges and tendencies. *Assessment and Evaluation in Higher Education*, 37(4), 453–464. <https://doi.org/https://doi.org/10.1080/02602938.2010.545868>
- López-Pastor, V. M., Pérez-Pueyo, Á., Barba, J. J., & Lorente-Catalán, E. (2016). Percepción del alumnado sobre la utilización de una escala graduada para la autoevaluación y coevaluación de trabajos escritos en la formación inicial del profesorado de educación física (FIPEF). *Cultura, Ciencia y Deporte*, 11(31), 37–50. <https://doi.org/10.12800/ccd.v11i31.641>
- López Pastor, V. M., Barba Martín, J. J., Monjas Aguado, R., Heras Bernardino, C., González Pascual, M., & Manrique Arribas, J. C. (2007). Trece años de evaluación compartida en Educación Física. *Revista Internacional de Medicina y Ciencias de La Actividad Física y Del Deporte*, 7(26), 69–86.
- López Pastor, V. M., González Pascual, M., & Barba Martín, J. J. (2005). La participación del alumnado en la evaluación: la autoevaluación, la coevaluación y la evaluación compartida. *Tándem: Didáctica de La Educación Física*, 17, 21–37.
- López Pastor, V. M., Monjas Aguado, R., Gómez García, J., López Pastor, E., Martín Pinela, J., González Badiola, J., Barba Martín, J. J., Aguilar Baeza, R., González Pascual, M., Heras Bernardino, C., Martín, M. I., Manrique Arribas, J. C., & Marugán García, L. (2015). La evaluación en educación física. Revisión de los modelos tradicionales y planteamiento de una alternativa: La Evaluación Formativa y Compartida. *Retos*, 2041(10), 31–40. <https://doi.org/10.47197/retos.v0i10.35061>
- López Pastor, V. M., & Pérez Pueyo, Á. (2017). *Evaluación formativa y compartida en educación: experiencias de éxito en todas las etapas educativas*. Universidad de León.
- Macken, S., MacPhail, A., & Calderon, A. (2020). Exploring primary pre-service teachers' use of 'assessment for learning' while teaching primary physical education during school placement. *Physical Education and Sport Pedagogy*, 25(5), 539–554. <https://doi.org/10.1080/17408989.2020.1752647>
- Martos-Garcia, D., Usabiaga, O., & Valencia-Peris, A. (2017). Students' perception on formative and shared assessment: Connecting two universities through the blogosphere. *Journal of New Approaches in Educational Research*, 6(1), 64–70. <https://doi.org/10.7821/naer.2017.1.194>
- McArthur, J. (2019). *La evaluación: una cuestión de justicia social Perspectiva crítica y prácticas adecuadas*. Narcea.
- Otero-Saborido. (2020). Evaluación formativa y compartida y aprendizaje invertido: narración de un caso en la enseñanza universitaria. In M. Reyes-Tejedor, D. Cobos-Sanchiz, & E. López-Meneses (Eds.), *Innovación Pedagógica Universitaria: Reflexiones y Estrategias* (pp. 175–188). Octaedro.
- Otero-Saborido, F. M., Sánchez-Oliver, A. J., Grimaldi-Puyana, M., & Álvarez-García, J. (2018). Flipped learning and formative evaluation in higher education. *Education and Training*. <https://doi.org/10.1108/ET-12-2017-0208>
- Pérez-Pueyo, Á., Barba-Martín, R., López-Pastor, V. M., & Lorente-Catalán, E. (2017). La utilización de las escalas graduadas de evaluación en la enseñanza universitaria. In Victor Manuel López-Pastor & Á. Pérez-Pueyo (Eds.), *Evaluación formativa y compartida en educación: experiencias de éxito en todas las etapas educativas* (pp. 374–384). Universidad de León. <https://buleria.unileon.es/handle/10612/5999>
- Pérez-Pueyo, Á., & López-Pastor, V. M. (2017). El estilo actitudinal como propuesta metodológica vinculada a la evaluación formativa. In V. M. López-Pastor & Á. Pérez-Pueyo (Eds.), *Evaluación formativa y compartida en educación: experiencias de éxito en todas las etapas educativas* (pp. 240–257). Universidad de León. <https://buleria.unileon.es/handle/10612/5999>
- Romero Martín, M. R., Asún Dieste, S., & Chivite Izco, M. T. (2015). La autoevaluación en expresión corporal en formación inicial del profesorado de educación física: un ejemplo de buena práctica (Self-assessment in body expression in Physical Education Teacher Education. An example of good practice). *Retos*, 2041(29), 236–241. <https://doi.org/10.47197/retos.v0i29.42493>
- Ruiz Martín, H. (2020). *Cómo aprendemos*. Graó.
- Sanchez, C. E., Atkinson, K. M., Koenka, A. C., Moshontz, H., & Cooper, H. (2017). Self-grading and peer-grading for formative and summative assessments in 3rd through 12th grade classrooms: A meta-analysis. *Journal of Educational Psychology*, 109(8), 1049–1066. <https://doi.org/10.1037/edu0000190>
- Santos-Pastor, M. L., & Martínez-Muñoz, L. F. (2017). El pacto en el aprendizaje desde la negociación de la rúbrica de evaluación. In *Evaluación formativa y compartida en educación: experiencias de éxito en todas las etapas educativas* (pp. 402–421). Universidad de León.
- Santos Guerra, M. Á. (2003). *La evaluación como aprendizaje: cuando la flecha impacta en la diana*. Madrid/ : Narcea, D.L. 2003.
- Soytürk, M. (2019). Analysis of Self and Peer Evaluation in Basic Volleyball Skills of Physical Education Teacher Candidates. *Journal of Education and Learning*, 8(2), 256. <https://doi.org/10.5539/jel.v8n2p256>
- Wisniewski, B., Zierer, K., & Hattie, J. (2020). The Power of Feedback Revisited: A Meta-Analysis of Educational Feedback Research. *Frontiers in Psychology*, 10. <https://doi.org/10.3389/fpsyg.2019.03087>