



**Universidad**  
Zaragoza

## Trabajo Fin de Máster

DIFERENCIAS ENTRE HETEROEVALUACIÓN,  
AUTOEVALUACIÓN Y COEVALUACIÓN EN  
EDUCACIÓN FÍSICA

DIFFERENCES BETWEEN HETEROEVALUATION,  
SELF-ASSESSMENT AND CO-ASSESSMENT IN  
PHYSICAL EDUCATION

*Autor/es*

David Garzo Sanagustín

*Director/es*

María Rosario Romero Martín

FACULTAD DE CIENCIAS HUMANAS Y DE LA EDUCACIÓN  
Año 2025

## ÍNDICE

RESÚMEN .....	1
ABSTRACT .....	2
1. INTRODUCCIÓN.....	3
1.1. Contextualización .....	3
1.2. Justificación de la elección.....	3
2. MARCO TEÓRICO .....	4
2.1. La evaluación formativa.....	4
2.2. La heteroevaluación .....	5
2.3. La autoevaluación.....	5
2.4. La coevaluación.....	6
3. OBJETIVOS .....	7
4. METODOLOGÍA.....	7
4.1. Diseño del estudio .....	7
4.2. Población de estudio .....	8
4.3. Instrumentos de investigación .....	9
4.4. Programa de intervención .....	11
4.5. Procedimiento.....	12
4.6. Análisis de datos.....	12
5. RESULTADOS .....	15
5.1. Análisis descriptivo .....	15
5.2. Análisis de asociaciones entre variables .....	17
5.3. Análisis de correlaciones .....	21
5.4. Análisis cualitativo .....	22
6. DISCUSIÓN .....	24
7. CONCLUSIONES.....	28
8. LIMITACIONES DEL ESTUDIO .....	29
9. REFERENCIAS BIBLIOGRÁFICAS.....	31
ANEXO I: TABLA ÍTEMS DEL CUESTIONARIO .....	35
ANEXO II. TABLAS ANÁLISIS CUALITATIVO.....	36

## RESÚMEN

Este trabajo analiza las percepciones del alumnado de Educación Secundaria sobre tres modalidades de evaluación formativa en la asignatura de Educación Física: heteroevaluación, autoevaluación y coevaluación.

El estudio se lleva a cabo en el I.E.S. Sierra de Guara (Huesca), con una muestra de 43 alumnos y alumnas de 4º de la ESO distribuidos en tres grupos, cada uno expuesto a una modalidad diferente. La recogida de datos se realiza mediante un cuestionario elaborado específicamente para este estudio, que incluye preguntas cerradas, abiertas y de escala Likert.

Los resultados muestran una preferencia mayoritaria por la evaluación realizada por el docente, aunque existe una apertura hacia modelos compartidos. La autoevaluación es considerada más útil que la coevaluación, aunque ambas son aceptadas como parte de la calificación final. El análisis cualitativo evidencia dificultades relacionadas con la objetividad, la falta de experiencia y la necesidad de formación en estas prácticas.

Finalmente, se destaca la importancia de formar al alumnado en competencias evaluativas y de establecer criterios comunes que permitan una aplicación eficaz y significativa de la evaluación formativa.

**Palabras clave:** Evaluación, Autoevaluación, Coevaluación, Educación Física, Secundaria

## ABSTRACT

This study analyzes the perceptions of secondary school students regarding three formative assessment modalities in Physical Education: hetero-assessment, self-assessment, and peer-assessment.

The research was conducted at I.E.S. Sierra de Guara (Huesca) with a sample of 43 4th-year ESO students divided into three groups, each exposed to a different assessment method. Data collection was carried out through a custom-designed questionnaire including closed-ended, open-ended, and Likert-scale questions.

Results show a predominant preference for teacher-led assessment, although there is openness to shared models. Self-assessment is perceived as more useful than peer-assessment, although both are generally accepted as part of the final grade. The qualitative analysis reveals difficulties related to objectivity, lack of experience, and the need for training in these practices.

The study underscores the importance of equipping students with evaluative competencies and establishing common criteria to ensure the effective and meaningful implementation of formative assessment.

**Keywords:** Formative assessment, Self-assessment, Peer-assessment, Physical Education, Secondary education.

## **1. INTRODUCCIÓN**

### **1.1. Contextualización**

La evaluación constituye un componente esencial en el proceso de enseñanza-aprendizaje de la Educación Física, permitiendo no solo valorar los resultados, sino también los procesos de adquisición de los diferentes conocimientos. En el ámbito de la Educación Física, cuyas características incluyen una mayor practicidad y dinamismo y donde la evaluación exige enfoques más integradores y continuos, la evaluación formativa es considerada una herramienta clave, orientada a la mejora del aprendizaje mediante un feedback constante, una reflexión y una participación activa por parte del alumnado.

Dentro del marco de la evaluación formativa, cobran especial importancia tres modalidades como lo son la heteroevaluación, la autoevaluación y la coevaluación. En primer lugar, la heteroevaluación implica una valoración del rendimiento del alumnado únicamente por parte del docente, permitiendo orientar el aprendizaje desde una perspectiva de un profesional de la educación y de la actividad física en este caso. La autoevaluación, por su parte, ofrece al estudiante la posibilidad de valorar su propio desempeño, priorizando la autorreflexión del propio alumno. Finalmente, la coevaluación fomenta la evaluación entre iguales, generando momentos de diálogo, colaboración y crítica constructiva.

Este trabajo se centra en explorar el papel de la evaluación Formativa en Educación Física, analizando las diferencias entre la heteroevaluación, autoevaluación y coevaluación aplicada a diferentes grupos de alumnos.

El presente estudio se realizará con estudiantes del 4º curso de Educación Secundaria Obligatoria del I.E.S. Sierra de Guara de Huesca, pertenecientes a tres grupos diferentes.

### **1.2. Justificación de la elección**

La elección de este tema responde a la relevancia que toma la evaluación formativa en el contexto actual de la educación, buscando un cambio metodológico hacia una enseñanza más competencial, participativa e inclusiva, tratando de dejar de lado los resultados físicos y priorizando otros como la actitud, la mejora individual o el esfuerzo.

Además, la inclusión de herramientas como la heteroevaluación, la autoevaluación y la coevaluación permite valorar el aprendizaje desde diferentes perspectivas, fomentando la autonomía

y la reflexión personal.

## **2. MARCO TEÓRICO**

### **2.1. La evaluación formativa**

La evaluación formativa se ha convertido en una estrategia pedagógica fundamental en el ámbito de la Educación Física, al permitir valorar el progreso del alumnado durante el proceso de enseñanza-aprendizaje, promoviendo su implicación activa, el desarrollo de competencias y la mejora continua (Bernardino et al., 2006). A diferencia de la evaluación tradicional centrada en la calificación, la evaluación formativa se enfoca en proporcionar información útil para ajustar las prácticas educativas, facilitar la reflexión y orientar los aprendizajes (Mamani Ramos, 2017).

Dentro de la especialidad de la Educación Física, que tiene un carácter procedimental y vivencial, la evaluación formativa se convierte en una de las alternativas más adecuadas. A través de esta, es posible valorar, a parte de los logros motrices, aspectos actitudinales, sociales y cognitivos, generando un enfoque integral del aprendizaje (Pérez Pueyo et al., 2008).

Numerosas investigaciones en este campo indican que una evaluación centrada en el feedback, la autoevaluación y la coevaluación fomenta la responsabilidad, la autonomía y la motivación del alumnado, así como una mayor conciencia de sus propios procesos de aprendizaje (Pérez Pino et al., 2017). Este planteamiento es especialmente útil en Educación Física ya que, por sus características, los resultados individuales no siempre son estandarizados y el desarrollo personal, la cooperación y la mejora continua deben tomar una gran importancia en la evaluación.

Como se destaca en el artículo de Pérez Pueyo et al. (2008), el estilo actitudinal es la metodología que vertebra el proceso educativo desde la evaluación formativa promoviendo que el alumnado analice sus errores, proponga mejoras y participe activamente en su propio aprendizaje. En este modelo, la evaluación deja de ser la parte final del proceso de enseñanza-aprendizaje para convertirse en un proceso continuo, compartido y significativo, que pone el foco en el por qué, para qué y para quién se evalúa, más allá del qué, cuándo o cómo.

Por otro lado, la implementación de la evaluación formativa en Educación Física por parte del profesorado todavía presenta ciertas dificultades, especialmente relacionadas con la planificación sistemática, la implicación del alumnado en el proceso y la evaluación de aspectos actitudinales (Cañadas & Santos-Pastor, 2021). No obstante, se confirma un avance hacia una evaluación más reflexiva y adaptada a las necesidades reales del alumnado.

## 2.2. La heteroevaluación

La heteroevaluación, entendida como el proceso de valoración y juicio que realiza el docente sobre el aprendizaje y desempeño del estudiante (Álvarez Méndez, 2002), ha sido tradicionalmente la modalidad evaluativa predominante en los sistemas educativos. En la materia de Educación Física, permite al docente observar, analizar y calificar cualitativa y cuantitativamente las competencias motrices, habilidades sociales, actitudes y conocimientos que el alumnado va adquiriendo a lo largo del proceso educativo.

La heteroevaluación no se limita a la asignación de una calificación numérica final, su verdadero valor está en la capacidad del docente para ofrecer un feedback constructivo (Vázquez Avilés, 2019). Este feedback debe ser específico, descriptivo y orientado a la mejora del alumnado, informando sobre sus aciertos y errores y guiándolo a la manera de superar sus dificultades.

La efectividad de la heteroevaluación reside en la claridad de los criterios de evaluación, tanto por parte de docente como de los estudiantes, quiénes deben conocer y comprender qué y cómo se va a evaluar (Fuentes Nieto, 2019). Además, la heteroevaluación debe ser diversa en cuanto a instrumentos y momentos de aplicación. Las rúbricas son el elemento más útil para este proceso, aunque se deben utilizar también listas de control, registros anecdóticos o pruebas de conocimientos conceptuales (Jiménez Barret, 2017).

## 2.3. La autoevaluación

La autoevaluación se define como la capacidad del estudiante para analizar, juzgar y valorar su propio aprendizaje, desempeño y progreso, basándose en criterios preestablecidos (Tirado González, 2009). Tal y como afirma en este mismo artículo Tirado Gonzalez (2009), la autoevaluación busca promover la metacognición, es decir, la reflexión crítica sobre el propio proceso de aprender, la identificación de fortalezas y áreas de mejora, y el desarrollo de estrategias para la autorregulación del aprendizaje.

En el contexto de la Educación Física, la autoevaluación es crucial para el desarrollo de la autonomía y la responsabilidad del alumnado. Cuando los estudiantes participan activamente en la valoración de su movimiento, su participación, su esfuerzo y su comprensión conceptual, no solo se apropian de su aprendizaje, sino que también desarrollan habilidades esenciales de pensamiento crítico y autorregulación (Cervantes & Medina, 2020).

Para que la autoevaluación sea efectiva, es indispensable que los estudiantes dispongan de

criterios de evaluación claros, comprensibles y conocidos por el alumno. Estos criterios, que deben ser compartidos y discutidos previamente, les sirven de referencia para realizar una valoración objetiva de su propio trabajo. Para el guiado de este proceso, es imprescindible la utilización de instrumentos como rúbricas de autoevaluación, listas de cotejo o cuestionarios reflexivos, adaptados al nivel y edad del alumnado, ya que la capacidad de autoevaluación va aumentando con la edad (Miras & Granda, 2018).

#### **2.4. La coevaluación**

La coevaluación, definida como el proceso de valoración mutua que realizan los estudiantes sobre el desempeño de sus pares, es un componente vital de la evaluación formativa (A. Cervantes & Medina, 2020). En el contexto de la Educación Física, la coevaluación no solo ofrece una perspectiva adicional sobre el aprendizaje, sino que también fomenta el desarrollo de habilidades sociales, comunicativas y cognitivas de gran valor para el alumnado.

Al igual que en los métodos anteriores, la claridad y conocimiento de los criterios de evaluación es un punto clave en la coevaluación ya que además, se evidencia que este proceso de coevaluación puede suponer desafíos al alumnado al inicio del mismo, sobre todo en la precisión de la evaluación entre pares (Alstot, 2018)

Además, la coevaluación promueve la responsabilidad y el trabajo colaborativo. Al evaluar a sus compañeros, los estudiantes se involucran activamente en el proceso de aprendizaje de los demás, ofreciendo feedback constructivo que puede ser percibido como más accesible y relevante por parte de quien la recibe, dado el contexto de igualdad entre pares (Backman et al., 2024). La coevaluación también expone a los estudiantes a diferentes enfoques para resolver problemas motrices, ampliando sus conocimientos. Para que este proceso sea efectivo, es fundamental establecer directrices claras y concisas sobre cómo se debe llevar a cabo la evaluación, qué aspectos deben valorarse y cómo debe ser el feedback (Cañadas & Santos-Pastor, 2021).

### **3. OBJETIVOS**

Una vez expuesto el marco teórico del presente Trabajo de Fin de Máster, se establecen los objetivos que se proponen para el estudio llevado a cabo.

Objetivos principales:

- Analizar el impacto de la aplicación de la heteroevaluación, autoevaluación y coevaluación en la opinión sobre la evaluación del alumnado de cuarto de la ESO del I.E.S. Sierra de Guara.
- Comparar las percepciones de aprendizaje en relación con las distintas modalidades de evaluación aplicadas.

Objetivos secundarios:

- Explorar las preferencias del alumnado acerca de las características de los diferentes modelos de evaluación a la hora de aplicarlos y la importancia de los resultados.
- Interpretar cualitativamente las sugerencias del alumnado sobre cómo mejorar su forma de evaluarse y evaluar a los demás.

En el estudio desarrollado en el presente TFM se pretende observar y analizar las percepciones del alumnado de Educación Secundaria en relación con los distintos tipos de evaluación formativa aplicados en las clases de Educación Física, la heteroevaluación, la autoevaluación y la coevaluación.

### **4. METODOLOGÍA**

#### **4.1. Diseño del estudio**

El presente estudio tiene un enfoque mixto al combinar metodologías cuantitativas y cualitativas, con el objetivo de obtener una comprensión más amplia de las percepciones del alumnado sobre la evaluación en la asignatura de Educación Física. Esta estrategia permite tanto el análisis estadístico de tendencias generales como la interpretación de opiniones individuales.

En primer lugar, se abordará el componente cuantitativo del estudio, en el cual se tratarán tres apartados diferenciados, el análisis de datos descriptivos, el análisis de asociaciones de variables y por último el análisis de correlaciones.

En el análisis de datos descriptivos se calcularán estadísticas descriptivas para las diferentes

variables del cuestionario, en este caso se hará un análisis de frecuencias dado que son variables nominales y por tanto no se pueden llevar a cabo pruebas como la media, la moda o la desviación típica. Estas pruebas proporcionan una visión general de las preferencias y percepciones dominantes del alumnado respecto a la evaluación en Educación Física (Rendón-Macías et al., 2016).

En el análisis de asociaciones entre variables se realizarán comparaciones entre los diferentes grupos encuestados en relación a las variables que se han considerado interesantes a estudiar, con el fin de identificar posibles diferencias significativas en las percepciones y preferencias de cada grupo.

Para finalizar el análisis cuantitativo, se llevará a cabo el análisis de correlaciones, donde se buscarán posibles interacciones entre variables del cuestionario para comprender mejor el nexo de unión entre el comportamiento de unas y otras variables. Algunas de las correlaciones estudiadas son la correlación entre la preferencia sobre quién evalúa y utilidad percibida de la auto/coevaluación, la correlación entre qué ayuda más a aprender y percepciones sobre evaluación, la correlación entre la aceptación del uso de la auto/coevaluación para la nota final y su utilidad percibida o la correlación entre la conciencia de progreso y aprendizaje asociado a la auto/coevaluación.

El análisis cualitativo se basará en el análisis de contenido de las preguntas abiertas incluidas en el cuestionario (Schettini & Cortazzo, 2016), que invitan al alumnado a reflexionar sobre cómo mejorar su forma de autoevaluarse y cómo mejorar su forma de evaluar a otros, permitiendo comprender de mejor manera las motivaciones y limitaciones del alumnado respecto a la evaluación, complementando así la estadística cuantitativa.

## **4.2. Población de estudio**

La población objeto de este estudio está compuesta por los grupos de 4º curso de Educación Secundaria Obligatoria del Instituto de Educación Secundaria Sierra de Guara, ubicado en Huesca. En concreto, han participado los estudiantes pertenecientes a tres grupos: 4ºA + 4ºDIV (grupo de diversificación), 4ºB y 4ºC.

Los criterios de inclusión fueron estar matriculado en alguno de los grupos mencionados, haber asistido a un mínimo del 75% de las sesiones impartidas y haber aceptado voluntariamente participar en este estudio, a través de la sección inicial del cuestionario, tal y como estipula el modelo de cuestionario proporcionado por la Universidad de Zaragoza. Por lo tanto, pese a que el

total de matrículas pertenecientes a estos grupos es de 57, la muestra final del estudio fue de 43 alumnos y alumnas de 4º de la ESO del I.E.S. Sierra de Guara.

En cuanto a las características de los grupos, existe una alta cantidad de alumnado absentista, a ello se debe la bajada en la muestra del estudio. Son grupos homogéneos en cuanto a la ratio de hombres y mujeres, siendo en 4ºA+DIV 9 mujeres y 5 hombres, en 4º B 8 mujeres y 5 hombres y en 4ºC 10 mujeres y 6 hombres, dejando el total de la muestra en 27 mujeres y 16 hombres.

**Tabla 1.** Recuento hombres y mujeres por grupo.

	4ºA+DIV	4º B	4ºC	TOTAL
Hombres	5	5	6	16
Mujeres	9	8	10	27

Desde el punto de vista motor, se observa que los grupos son homogéneos, con niveles de habilidad variados, desde alumnado con una motricidad desarrollada hasta otros con dificultades en coordinación y percepción espacio-temporal. A nivel de relaciones sociales, la mayoría del alumnado muestra habilidades sociales adecuadas, aunque existen casos puntuales de alumnos con menor participación. En cuanto a la atención y la implicación, la mayor parte del grupo se mantiene activo y con buena disposición hacia las tareas, aunque hay alumnado que pierde el interés por el juego, sobre todo en los últimos tramos de las sesiones, debido en parte a la larga duración de las sesiones. En lo que respecta al aprendizaje, las dificultades se centran en la comprensión táctica del juego y la toma de decisiones en situaciones de juego real. Estas características fueron extraídas mediante una observación previa del grupo durante el periodo de prácticas, así como de la entrevista con el tutor del centro.

### 4.3. Instrumentos de investigación

Para la recogida de datos se ha utilizado un cuestionario, diseñado específicamente para este estudio. El instrumento cuenta con 15 preguntas entre las que encontramos preguntas de respuesta cerrada, de respuesta abierta y de escala Likert (cinco opciones de respuesta que van desde nunca, casi nunca, a veces, casi siempre y siempre), divididas en los tres bloques siguientes:

- Cuestiones generales: incluye ítems relacionados con las preferencias del alumnado

respecto a quién y cuándo debe evaluarle, así como los métodos de evaluación que consideran más útiles para aprender.

- Autoevaluación: recoge la opinión del alumnado sobre la utilidad de evaluarse a sí mismos, las dificultades que encuentran en el proceso, su percepción de justicia al incluir la autoevaluación en la calificación final y su grado de competencia para autoanalizar su propio rendimiento. Incluye además una pregunta abierta para profundizar en cómo mejorar esta práctica.

- Coevaluación: cuestiona la percepción del alumnado sobre la evaluación entre compañeros, su capacidad para valorar de forma objetiva el trabajo de otros, la utilidad que le atribuyen a este tipo de evaluación y su opinión sobre su inclusión en la nota final. También se incorpora una pregunta abierta sobre posibles mejoras en su forma de evaluar a otros.

Para el diseño del cuestionario, se han tenido en cuenta la estructura y contenido de instrumentos empleados por otros autores en estudios previos sobre la evaluación formativa en Educación Física (López-Pastor et al., 2016). Además, para darle coherencia con el contexto del estudio y el marco teórico del mismo, se consideraron algunos aspectos de Cuestionarios Validados sobre el feedback entre iguales en educación superior (Sánchez-Martínez et al., 2019).

Con el fin de garantizar la validez del contenido del cuestionario y la correcta comprensión de los ítems del mismo, se utilizó la técnica de validez del contenido por juicio de expertos (Galicia Alarcón et al., 2017), para lo que se contactó con tres profesores universitarios expertos en el área de la Educación Física, expertos en evaluación formativa a partir de su cálculo de experticia obtenido teniendo en cuenta tres criterios: 1. Publicaciones en evaluación; 2. Participación en proyectos de investigación del tema y 3. Docencia en didáctica, Educación Física y evaluación formativa, a los que se les envió el cuestionario y se les pedía que informasen sobre: 1. Claridad del contenido; 2. Pertinencia de los ítems descritos y 3. Completitud del cuestionario, con el fin de corroborar si el conjunto del cuestionario recogía la información necesaria para responder con fiabilidad el objeto del estudio.

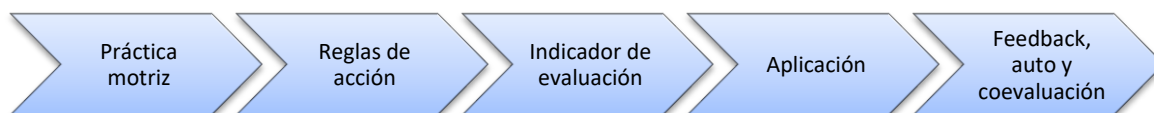
Tras esta consulta, el cuestionario se vio modificado teniendo en cuenta cada una de las sugerencias propuestas, como por ejemplo la redacción de algunos ítems, la fusión de dos de ellos y la ampliación de otro, dando por resultado el cuestionario definitivo que sería proporcionado a los

alumnos de forma anónima y voluntaria a través de un Código QR durante la última sesión impartida, que les direccionó directamente al cuestionario creado con la herramienta Google Forms, conforme a las indicaciones éticas recogidas por la Universidad de Zaragoza y la normativa vigente en materia de protección de datos.

#### 4.4. Programa de intervención

La intervención se llevó a cabo durante el periodo de prácticas del Prácticum II, en el IES Sierra de Guara (Huesca), concretamente con los grupos de 4º de ESO A+DIV, 4ºB y 4ºC, durante el desarrollo de una Unidad Didáctica de Bádminton de cuatro sesiones dobles, que conformarían ocho sesiones al uso. Durante esta unidad se trabajó con el método comprensivo, centrándose en la táctica del bádminton. En lo que respecta a este estudio, el propósito de estas sesiones fue aplicar distintas modalidades de evaluación según el agente evaluador, de manera que cada grupo experimentó una forma específica de evaluación a lo largo de las sesiones. En concreto, el grupo 4ºA+DIV trabajó con heteroevaluación, el grupo 4ºB con autoevaluación, y el grupo 4ºC con coevaluación. Esta distribución se hizo de manera aleatoria y permitió establecer una variable independiente clara para el análisis posterior: la modalidad de evaluación vivenciada.

La metodología empleada durante la unidad se basó en el enfoque del descubrimiento guiado, en base a lo propuesto por Estrada Marcén y Romero Martín (2024), partiendo de situaciones de aprendizaje prácticas en las que el alumnado, a través de la experimentación y la reflexión dirigida, era conducido a establecer reglas de acción relacionadas con el juego y la táctica. A partir de estas reglas, el propio alumnado identificaba, con ayuda del docente, al menos un indicador de evaluación, el cual servía como criterio observable para valorar el rendimiento. Este indicador era posteriormente aplicado en la última actividad de cada sesión en cada grupo en función de su modalidad correspondiente a través de unas rúbricas de un único ítem (Estrada Marcén & Romero Martín, 2024). Estas rúbricas estaban compuestas por cinco indicadores de logro (Excelente, Muy bueno, Bueno, Aceptable y Mejorable), junto con la correspondiente descripción que se ajusta a cada indicador de logro.



**Figura 1.** Esquema del proceso metodológico llevado a cabo

#### 4.5. Procedimiento

La investigación comenzó con la preparación de las sesiones, orientadas a los objetivos que se querían evaluar en las rúbricas, así como el diseño y la redacción de estas. Posteriormente, se llevó a cabo la Unidad Didáctica en el centro, aplicando las rúbricas y la metodología comentada en el anterior apartado.

Al finalizar la última sesión del Prácticum II, se les pasó a través de un Código *QR* el formulario previamente descrito, con el fin de recoger las percepciones y opiniones del alumnado sobre la evaluación llevada a cabo. Este cuestionario fue rellenado por todo aquel alumnado que aceptó participar en él en ese mismo instante, aceptando las condiciones que establece la plantilla de la estructura adecuada de un cuestionario, propuesta por la Universidad de Zaragoza. Además, para aquellos alumnos que no pudieron asistir a esta última sesión, se les envió el cuestionario a través de Google Classroom para que pudieran cumplimentarlo en otro momento.

#### 4.6. Análisis de datos

Una vez finalizada la recogida de datos, las respuestas obtenidas a través del cuestionario fueron exportadas a una hoja de cálculo en Microsoft Excel. Para su correcto tratamiento estadístico, se procedió a la conversión de las respuestas a valores numéricos, permitiendo su análisis cuantitativo. Posteriormente, estos datos fueron exportados al programa estadístico PSPP, donde se continuó con el análisis cuantitativo.

Con estos datos ya importados en PSPP, se asignaron etiquetas de valor a cada una de las variables codificadas, de forma que cada número reflejara su correspondiente respuesta original del cuestionario, con el fin de que los resultados se visualizaran correctamente en las tablas obtenidas del análisis estadístico. Finalmente, el archivo se dividió en los tres grupos existentes para obtener una tabla de cada uno de los grupos y poder realizar las comparaciones necesarias, tal y como indica el procedimiento correcto del programa PSPP (Araujo, 2014).

El análisis de datos constará de dos bloques, el bloque cuantitativo, que a su vez se divide en análisis descriptivo, análisis comparativo y análisis de correlaciones y el bloque cualitativo.

Ya con la base de datos estructurada, en primer lugar se aplicó la prueba de normalidad Shapiro-Wilk, indicada para muestras menores de cincuenta valores (Razali & Wah, 2011), sobre las variables cuantitativas, con el fin de concretar las pruebas que deberán utilizarse posteriormente para tratar los datos. Tras pasar esta prueba, se obtuvo un p-valor menor que 0,05 lo que permitió

determinar que los datos no seguían una distribución normal, tal y como se indica en el artículo de Razaly y Wah (2011). Por esta razón, se optó por utilizar pruebas no paramétricas para el análisis cuantitativo.

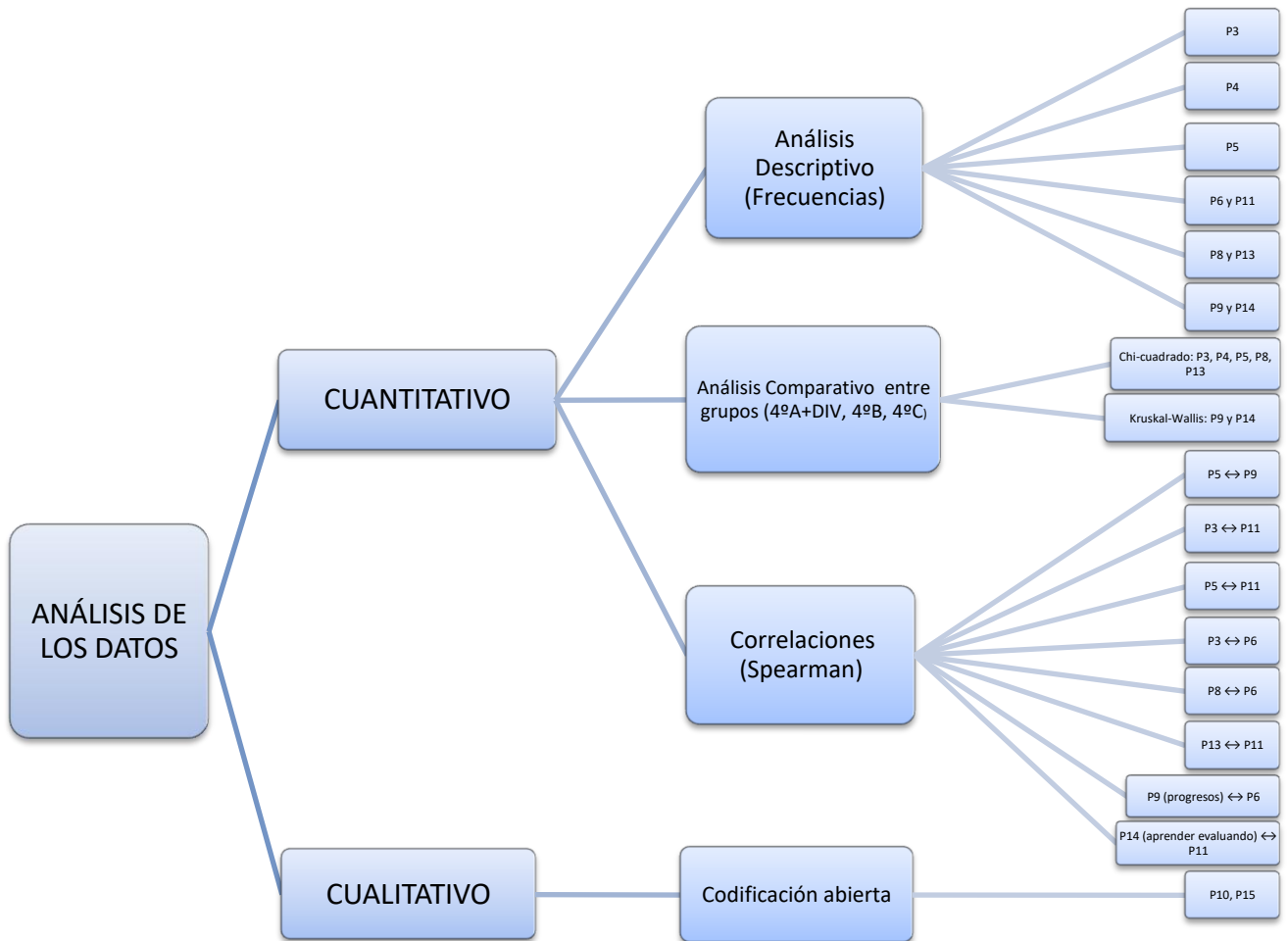
Comenzando por el primer bloque, el cuantitativo, en primer lugar, se realizó el análisis descriptivo, donde se calcularon frecuencias y porcentajes de los ítems escogidos, dejando de lado las escalas tipo Likert y las respuestas cualitativas. Estos ítems por tanto son la preferencia sobre quién debe evaluar (P3), el momento preferido de la evaluación (P4), aquello que ayuda más a aprender (P5), la utilidad percibida de la autoevaluación y la coevaluación (P6 y P11) y la aceptación de la utilización de la autoevaluación/coevaluación para la nota final (P8 y P13).

Posteriormente, se llevo a cabo un estudio de las posibles relaciones entre determinadas variables, y los valores de las mismas en los distintos grupos de clase, analizando las asociaciones cuando las variables eran categóricas. Para ello se utilizaron las pruebas Chi-cuadrado para las variables categóricas de P3, P4, P5, P8 Y P13, ya que esta prueba es adecuada para determinar si existe una asociación entre variables categóricas, además no requiere una distribución normal de los datos y es más adecuada cuando los datos se expresan en frecuencias (Amaya Cedrón, 2017). Para los ítems tipo Likert (P9 y P14) se utilizó la prueba de Kruksal-Wallis, que es una prueba no paramétrica para comparar las medias en variables ordinales, como alternativa a la prueba ANOVA de un factor (McKight & Najab, 2010). Además, para estas últimas preguntas, se calculó la media a través de lo que propone en su artículo Gutiérrez García et al. (2011) donde la que las variables ordinales pueden tratarse como métricas siempre y cuando estas tengan cinco o más categorías, permitiendo así operar con ellas.

Para terminar con el análisis cuantitativo, se realizó un análisis de correlaciones entre las siguientes preguntas del cuestionario: P3 con P6, P3 con P11, P5 con P9, P5 con P11, P8 con P6, P13 con P11, P9 (progresos) con P6 y P14 (aprender evaluando) con P11. Para el análisis de correlaciones se utilizó el coeficiente de correlación de Spearman, al tratarse de una prueba no paramétrica, para variables no normales (Mendivelso, 2021).

Por último, se continuó con el segundo bloque, el análisis cualitativo. Para ello, utilizó una metodología cualitativa para analizar las respuestas abiertas del cuestionario (P10 y P15), a partir de ello, se procedió a una codificación abierta de las respuestas, identificando categorías emergentes que muestran las opiniones y percepciones del alumnado respecto a la autoevaluación y coevaluación (Guillén Correas, 2014). Con este proceso se consiguió obtener una comprensión

contextualizada, identificando temas recurrentes sobre las debilidades que encuentran los alumnos en estos tipos de evaluación, ofreciendo así un enfoque más integral del estudio.



**Figura 2.** Esquema del análisis de datos

## 5. RESULTADOS

Una vez finalizada la recogida y tratamiento de datos, en este apartado se examinarán de forma detallada los resultados obtenidos en el cuestionario, no solo describiendo las preferencias de evaluación del alumnado participante, sino también comparando cada uno de los grupos examinados o explorando las relaciones entre diferentes variables. Además, se realizará también un análisis cualitativo sobre la percepción de los alumnos de sus puntos de mejora en la autoevaluación y coevaluación.

### 5.1. Análisis descriptivo

En primer lugar, se lleva a cabo el análisis descriptivo de ítems clave del cuestionario como los descritos a continuación.

Comenzando por la preferencia del alumnado sobre quién debe ser el agente evaluador (Tabla 2), observamos cómo la opción más elegida por el alumnado, con un 44,2% es que sea únicamente el docente quien evalúe, siendo a su vez la combinación de profesor, alumno y compañeros, con un 23,3% la opción de evaluación compartida más votada.

**Tabla 2.** Preferencia del agente evaluador.

	<b>Frecuencia</b>	<b>Porcentaje</b>
<b>Solo el profesor/a</b>	19	44,2%
<b>Solo el propio alumno/a</b>	2	4,7%
<b>El profesor/a y el alumno/a</b>	7	16,3%
<b>El profesor/a y mis compañeros/as</b>	5	11,6%
<b>Todos juntos</b>	10	23,3%
<b>Total</b>	43	100,0%

El siguiente ítem analizado es el momento de evaluación. Como se puede apreciar en la Tabla 3, la gran mayoría del alumnado, concretamente el 72,1%, prefiere que la evaluación sea continua a lo largo de toda la Unidad Didáctica, seguida esta por la opción de al terminar la Unidad Didáctica, con un porcentaje mucho menor, del 18,6% y finalmente, la opción menos preferida por los alumnos es que la evaluación se realice al comenzar la Unidad Didáctica, siendo únicamente el 9,3% de ellos quién ha escogido esta opción.

**Tabla 3.** Preferencia del momento de evaluación.

	<b>Frecuencia</b>	<b>Porcentaje</b>
<b>Al terminar la Unidad Didáctica</b>	8	18,6%
<b>Al inicio y al final de la Unidad didáctica</b>	4	9,3%
<b>A lo largo de todas las sesiones</b>	31	72,1%

En la Tabla 4, se muestran los elementos que el alumnado consideran que más les ayudan a aprender. Los comentarios del docente (60,5%) y la nota numérica (30,2%) se sitúan como las opciones más elegidas. Destaca el bajo porcentaje de las opciones que implican el trabajo con el resto de compañeros, siendo la suma de estos porcentajes de únicamente el 9,3% del total.

**Tabla 4.** Elemento que más ayuda a aprender.

	<b>Frecuencia</b>	<b>Porcentaje</b>
<b>Una nota numérica</b>	13	30,2%
<b>Comentarios del docente</b>	26	60,5%
<b>Comentarios de los compañeros</b>	3	7,0%
<b>Comparar mi trabajo con otros</b>	1	2,3%

A continuación, se exponen las tablas de utilidad tanto de la autoevaluación como de la coevaluación (Tabla 5). Como podemos observar, la mayoría del alumnado considera que ambas son útiles siempre o a veces, aunque en el caso de la autoevaluación se considera que siempre es útil en mayor frecuencia que la coevaluación, con un 60,5% frente al 30,2% de la coevaluación. Al mismo tiempo, un 7% de los alumnos consideran que la autoevaluación no es útil nunca, mientras que ninguno de ellos lo considera así de la coevaluación.

**Tabla 5.** Utilidad autoevaluación y coevaluación.

	<b>Autoevaluación</b>		<b>Coevaluación</b>	
	<b>Frecuencia</b>	<b>Porcentaje</b>	<b>Frecuencia</b>	<b>Porcentaje</b>
<b>Sí, siempre</b>	26	60,5%	13	30,2%
<b>A veces</b>	14	32,6%	6	14,0%
<b>No, nunca</b>	3	7,0%	24	55,8%

Para finalizar el análisis descriptivo, se estudia cómo los alumnos consideran que la autoevaluación y la coevaluación deben utilizarse para determinar la nota final de la asignatura. Tal y como nos muestran estos datos (Tabla 6), la gran mayoría creen que estos modelos de evaluación deben tenerse en cuenta en todas las ocasiones con un 55,8% en ambos casos y solo si coincide con la opinión del docente en un 37,2% para la autoevaluación y un 34,9% para la coevaluación. Esto indica por tanto que únicamente el 7% y el 9,3% en cada uno de los casos considera que la autoevaluación y la coevaluación no deben tomar parte en la nota de la asignatura.

**Tabla 6.** Utilización de la autoevaluación y coevaluación para la nota.

	Autoevaluación		Coevaluación	
	Frecuencia	Porcentaje	Frecuencia	Porcentaje
Sí	24	55,8%	24	55,8%
Solo si coincide con su opinión	16	37,2%	15	34,9%
No, porque uno puede mentir	1	2,3%	4	9,3%
No, la nota debe ponerla solo el profesor/a	2	4,7%	0	0%

## 5.2. Análisis de asociaciones entre variables

A continuación, se llevará a cabo el análisis de Asociaciones entre variables categóricas, con el fin de encontrar si los valores de alguna de ellas indican que los grupos de clase se comportan de manera diferente. Para ello, como ya se ha comentado previamente, se utiliza la prueba de Chi-cuadrado para las variables categóricas y la de Kruskal-Wallis para las de escala Likert.

En primer lugar, se presenta la prueba Chi-cuadrado aplicada a la variable del agente evaluador (Tabla 7). El valor de  $p=0,055$  por lo que hay que asumir la hipótesis nula ( $H_0$ ) que significa que no puede asociarse la variable de estudio al tipo de grupo clase, es decir, los tres grupos se comportan de manera semejante en cuanto a esta variable. No obstante, a nivel didáctico, puede ser interesante, tal y como se puede observar a través de las frecuencias, que el grupo 4°C prefiere una evaluación realizada únicamente por el profesor, mientras que los otros grupos muestran una distribución más variada.

**Tabla 7.** Comparación del agente evaluador por grupos

		Grupos			Total		Chi-cuadrado de Pearson	Razón de Semejanza
		4°A+DIV	4°B	4°C				
Quién evalúa	Solo el profesor/a	5	4	10	19	Valor	15,23	16,59
	Solo el propio alumno/a	0	1	1	2	df	8	8
	El profesor/a y el alumno/a	2	4	1	7	Sign, Asintótica (2-colas)	,055	,035
	El profesor/a y mis compañeros/as	0	3	2	5			
	Todos juntos	7	1	2	10			

Siguiendo con el análisis comparativo, en la Tabla 8 encontramos la comparación del momento de evaluación por grupos. En este caso, la  $p=0,071$  por lo que no existe significación y consideramos la hipótesis nula ( $H_0$ ) y por tanto no existen relaciones entre las respuestas de los diferentes grupos ya que todos responden de manera semejante. De igual forma, a nivel didáctico, observando las frecuencias, se aprecia como la gran mayoría de alumnos optan por una evaluación continua a lo largo de todas las sesiones, siendo en este caso 4°C la que menos opta por esta opción.

**Tabla 8.** Comparación momento de evaluación por grupos

		Grupos			Total		Chi-cuadrado de Pearson	Razón de Semejanza
		4°A+DIV	4°B	4°C				
Cuándo evalúa	Al terminar la Unidad Didáctica	3	0	5	8	Valor	8,63	11,83
	Al inicio y al final de la Unidad didáctica	0	1	3	4	df	4	4
	A lo largo de todas las sesiones	11	12	8	31	Sign, Asintótica (2-colas)	,071	,019

A continuación, en la Tabla 9 encontramos las diferencias entre grupos sobre el elemento que más ayuda a aprender, siendo esta comparación no significativa ( $p=0,413$ ), considerando por tanto la hipótesis nula ( $H_0$ ), que indica que no hay diferencias entre los grupos. La opción más votada para todos los grupos es la de los comentarios del docente.

**Tabla 9.** Comparación elemento que más ayuda a aprender por grupos

		Grupos			Total		Chi-cuadrado de Pearson	Razón de Semejanza
		4°A+DIV	4°B	4°C				
Ayuda a aprender	Una nota numérica	6	3	4	13	Valor	6,09	7,02
	Comentarios del docente	6	10	10	26	df	6	6
	Comentarios de los compañeros	1	0	2	3	Sign, Asintótica (2-colas)	,413	,319
	Comparar mi trabajo con otros	1	0	0	1			

El análisis de la utilización de la autoevaluación no muestra diferencias significativas entre los grupos ( $p=0,134$ ), en general en todos los grupos se decantan por la opción de que esta sí debe influir en la nota, ya sea coincidiendo o no con la opinión del profesor.

**Tabla 10.** Comparación utilización de la autoevaluación para la nota

		Grupos			Total		Chi-cuadrado de Pearson	Razón de Semejanza
		4ºA+DIV	4ºB	4ºC				
AUTOev para la nota,	Sí	5	7	12	24	Valor	9,79	10,29
	Solo si coincide con su opinión	7	5	4	16	df	6	6
	No, porque uno puede mentir	0	1	0	1	Sign, Asintótica (2-colas)	,134	,113
	No, la nota debe ponerla solo el profesor/a	2	0	0	2			

De la misma forma, la utilización de la coevaluación para la nota no muestra diferencias significativas entre los grupos ( $p=0,456$ ), ya que al igual que para la autoevaluación, la mayoría optan por tenerla en cuenta para la nota de la asignatura.

**Tabla 11.** Comparación utilización de la coevaluación para la nota

		Grupos			Total		Chi-cuadrado de Pearson	Razón de Semejanza
		4ºA+DIV	4ºB	4ºC				
Coev para la nota,	Sí	7	7	10	24	Valor	3,65	4,79
	Solo si coincide con su opinión	7	4	4	15	df	4	4
	No, porque uno puede mentir	0	2	2	4	Sign, Asintótica (2-colas)	,456	,310

Al analizar las opiniones sobre la autoevaluación por grupos (Tabla 12), observamos que en ninguna de las categorías existen diferencias significativas entre los grupos. Como observamos, las medias oscilan entre el 3,74 y el 4,14, lo que equivaldría a la respuesta “Casi siempre”.

**Tabla 12.** Comparación opinión sobre la autoevaluación por grupos

		N	Media	Chi-cuadrado	df	Sign, Asint,
Capacidad de evaluar sin ayuda externa	4ºA+DIV	14	3,36	5,94	2	,051
	4ºB	13	4,23			
	4ºC	16	3,69			
	Total	43	3,74			
Facilidad para entender los objetivos	4ºA+DIV	14	3,93	,94	2	,625
	4ºB	13	4,31			
	4ºC	16	4,19			
	Total	43	4,14			
Capacidad de identificar errores propios	4ºA+DIV	14	3,86	1,21	2	,547
	4ºB	13	4,15			
	4ºC	16	4,19			
	Total	43	4,07			
Autoevaluación ayuda a ser consciente de los progresos	4ºA+DIV	14	3,64	2,88	2	,237
	4ºB	13	4,23			
	4ºC	16	4,00			
	Total	43	3,95			

Algo similar ocurre en las respuestas de la opinión sobre la coevaluación por grupos (Tabla 13), donde no se observan diferencias significativas en ninguna de las variables. En este caso, destaca la media de la variable “Dificultad para ser objetivo en la coevaluación” que es de 3,02, equivalente a la respuesta “A veces”, lo que nos indica un mayor grado de dificultad para el alumnado en este aspecto respecto a las demás variables estudiadas cuyas medias oscilan entre 3,93 y 4,21.

**Tabla 13.** Comparación opinión sobre la coevaluación por grupos

		N	Media	Chi-cuadrado	df	Sign, Asint,
Capacidad de observar y valorar a un compañero	4ºA+DIV	14	3,71	5,34	2	0,069
	4ºB	13	4,62			
	4ºC	16	4,25			
	Total	43	4,19			
Facilidad para dar una opinión al compañero	4ºA+DIV	14	3,36	7,77	2	0,21
	4ºB	13	4,15			
	4ºC	16	4,25			
	Total	43	3,93			
Conciencia de aprendizaje en la coevaluación	4ºA+DIV	14	4,14	,66	2	0,719
	4ºB	13	4,23			
	4ºC	16	4,25			
	Total	43	4,21			
Dificultad para ser objetivo en la coevaluación	4ºA+DIV	14	2,86	4,75	2	0,093
	4ºB	13	2,54			
	4ºC	16	3,56			
	Total	43	3,02			

### 5.3. Análisis de correlaciones

Para finalizar con el análisis cuantitativo, se expone a continuación el análisis correlacional. Para llevar a cabo este, debido a las características de la prueba, se utilizó el coeficiente de correlación de Spearman.

En primer lugar, se analiza la correlación entre el agente evaluador y la utilidad de la autoevaluación y la coevaluación (Tabla 14), obteniendo como resultado que estas no tienen correlaciones significativas, con  $p=0,246$  y  $p=0,604$  respectivamente. Por otro lado, vemos que la correlación obtenida es negativa y baja en ambos casos.

**Tabla 13.** Correlación entre el agente de evaluación y utilidad de la autoevaluación y la coevaluación.

		Utilidad AUTOev,	Utilidad Coev,
Agente evaluador	Correlación de Spearman	-,181	-,081
	Sign, (2-colas)	,246	,604
	N	43	43

En la Tabla 15 encontramos la correlación entre el elemento que más ayuda a aprender y la utilidad de la autoevaluación y la coevaluación, donde no se observan correlaciones significativas entre estas variables. En el caso de la autoevaluación se aprecia una baja correlación negativa mientras que en la coevaluación existe una leve correlación positiva entre el elemento que más ayuda a aprender (comentarios del docente) y la utilidad de la coevaluación.

**Tabla 15.** Correlación entre qué ayuda a aprender y utilidad de la autoevaluación y la coevaluación.

		Utilidad AUTOev,	Utilidad Coev,
Ayuda a aprender	Correlación de Spearman	-,073	,240
	Sign, (2-colas)	,643	,121
	N	43	43

La Tabla 16 muestra la correlación entre la utilización de la autoevaluación para la nota y la utilidad de esta. La significación obtenida es levemente positiva aunque no significativa ( $p=0,61$ ).

**Tabla 16.** Correlación entre utilización de la autoevaluación para la nota y utilidad de la autoevaluación.

		Utilidad AUTOev,
AUTOev, para la nota	Correlación de Spearman	,288
	Sign, (2-colas)	,061
	N	43

De la misma forma que en la autoevaluación, la utilización de la coevaluación para la nota final (Tabla 17) tiene una correlación levemente positiva aunque no significativa con la utilidad de la misma, siendo en este caso  $p=0,193$ .

**Tabla 17.** Correlación entre utilización de la coevaluación para la nota y utilidad de la coevaluación.

		<b>Utilidad Coev,</b>
<b>COev, para la nota</b>	<b>Correlación de Spearman</b>	,202
	<b>Sign, (2-colas)</b>	,193
	<b>N</b>	43

La correlación entre la ayuda al progreso de la autoevaluación y la utilidad de la misma es la única correlación significativa ( $p=0,022$ ), sin embargo, esta correlación es negativa, lo que supone que quienes consideran que la autoevaluación ayuda a ser consciente del progreso, la califican como menos útil.

**Tabla 18.** Correlación entre ayuda al progreso de la autoevaluación y la utilidad de la autoevaluación.

		<b>Utilidad AUTOev,</b>
<b>Ayuda Progreso AUTOev,</b>	<b>Correlación de Spearman</b>	-,349 <sup>a</sup>
	<b>Sign, (2-colas)</b>	,022
	<b>N</b>	43

Por último, la Tabla 19 muestra la correlación entre aprender evaluando con la coevaluación y la utilidad de esta. Obtenemos una correlación negativa que además no es significativa. ( $p=0,256$ ).

**Tabla 19.** Correlación entre aprender evaluando en la coevaluación y la utilidad de la coevaluación.

		<b>Utilidad Coev,</b>
<b>Aprender evaluando COev,</b>	<b>Correlación de Spearman</b>	-,177
	<b>Sign, (2-colas)</b>	,256
	<b>N</b>	43

#### 5.4. Análisis cualitativo

El análisis cualitativo se va a llevar a cabo con aquellas preguntas del cuestionario que permitían respuesta abierta, P10 y P15 en este caso, que corresponden a las opiniones sobre cómo mejorar la forma de autoevaluar y coevaluar respectivamente. Para ello, se ha utilizado el enfoque que propone Guillen Correas (2017) en su libro, a través de una codificación abierta de variables y una clasificación categórica de estas (Anexo II).

Comenzando con la primera pregunta, relacionada con la mejora de la autoevaluación, se observaron varias categorías de pensamiento. Un gran número de alumnos consideraba que debían

ser más objetivos y críticos consigo mismos, con frases como “ser más objetivo” (122), “ser más objetiva conmigo misma y aceptar las críticas” (111) o “ser más objetivo y rectificar mis errores” (324). De igual forma, otro gran número de respuestas estaban relacionadas con el reconocimiento de los errores, que queda reflejado en respuestas como “aceptar errores” (325), “darme cuenta de mis errores antes” (319) o “fijarme más en mis errores” (322).

Otros grupos de alumnos, aunque en menor medida, destacaron la importancia de buscar un apoyo externo (“pedir consejo al profesor” (341), “pedir ayuda a un compañero” (335)), la falta de experiencia en la autoevaluación (“evaluarte más” (321), “practicar más” (316)) o utilizar herramientas para el análisis (“grabarme” (334)). Por último, también es destacable la cantidad de alumnos que no reflexionaron en esta pregunta, dejando respuestas como “no sé” (114), “nada” (328) o incluso respondiendo únicamente con un punto (109).

Por otro lado, analizando la segunda pregunta, en este caso centrada en la coevaluación, de la misma forma que en la primera, la respuesta más frecuente es la relacionada con la objetividad, destacando frases como “ser más objetiva sin tener en cuenta las amistades” (216) o “ser más objetiva, aunque sean mis compañeros o amigos tengo que evaluar tanto las partes buenas como las malas” (210). Otra categoría de respuestas que se repite en varias ocasiones es la relacionada con la concentración a la hora de evaluar, que queda reflejado en respuestas del tipo “fijarme más” (203), “estar más atento a los demás” (219) o “prestar mucha atención a cómo juegan” (243).

Algunas respuestas, aunque menos frecuentes, estaban relacionadas con el ayudar a mejorar al compañero durante el proceso (“ayudarles” (204), “darles consejos para que puedan hacerlo lo mejor posible” (245)) y a reconocer el esfuerzo de este (“ser consciente del esfuerzo de cada persona” (212), “valorar el esfuerzo de otro compañero” (214)). De la misma forma que en la autoevaluación, también se registraron respuestas vacías con nula reflexión, con ejemplos como “nada” (220), “no sé” (233) o “creo que ya lo hago bien” (238).

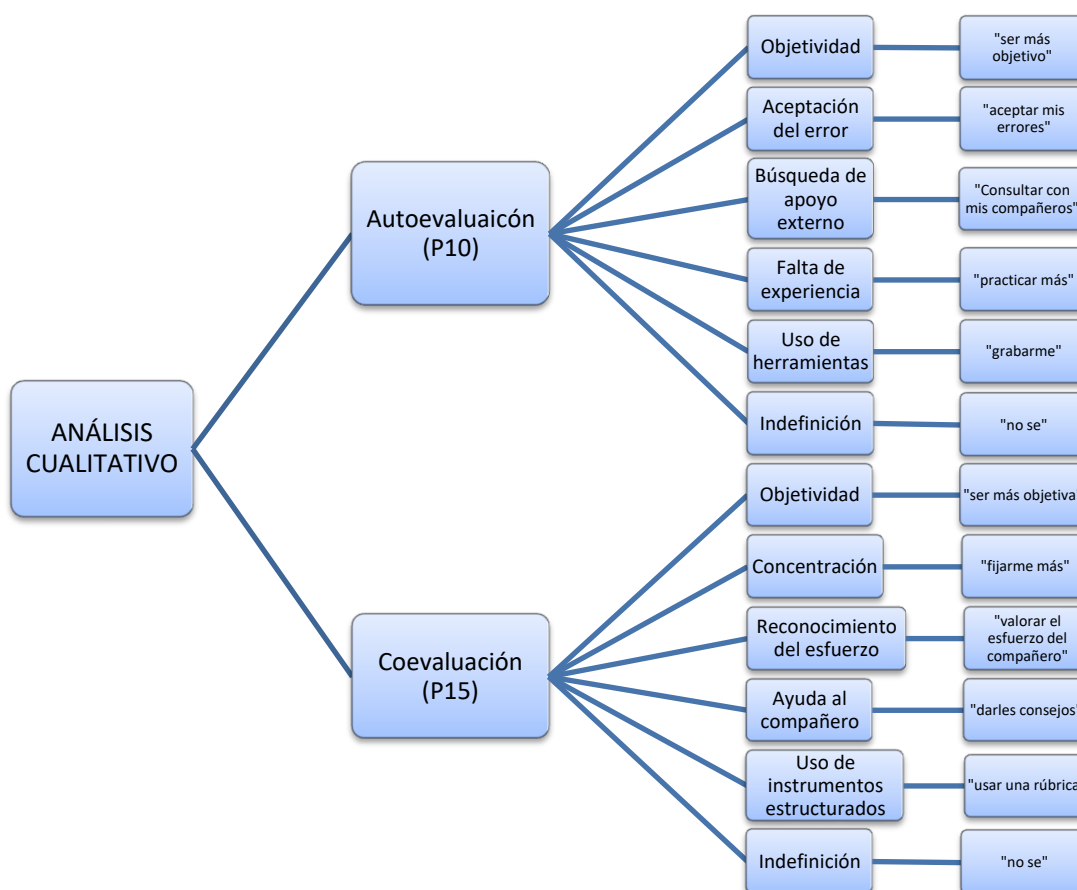


Figura 3. Categorías análisis cualitativo

## 6. DISCUSIÓN

A continuación, se va a proceder con la discusión de los resultados obtenidos en el estudio.

Desde el enfoque descriptivo, se evidencia que el alumnado muestra una fuerte preferencia por que la evaluación la realice únicamente el profesor, con el 44,2% de los estudiantes eligiendo esta opción (Tabla 1). Dato que resulta coherente con estudios que señalan que en contextos donde la evaluación tradicional ha sido la predominante durante muchos años, los alumnos tienden a asociar la figura del docente como referente de objetividad y justicia (Fuentes Nieto, 2019). No obstante, destaca también el 23,3% del alumnado que opta por una evaluación compartida entre el profesor, los alumnos y los compañeros, lo que confirma un cambio progresivo hacia modelos más participativos (Bernardino et al., 2006).

En cuanto al momento preferido para ser evaluados (Tabla 3), un 72,1% del alumnado se inclinó por una evaluación continua a lo largo de la Unidad Didáctica, lo que respalda lo propuesto en modelos donde la evaluación no es un fin en sí mismo, sino una herramienta de constante mejora

(Pérez Pueyo et al., 2008). Este tipo de evaluación permite una observación continua a lo largo del tiempo de los procesos de aprendizaje, promoviendo así una enseñanza significativa (Mamani Ramos, 2017).

Por su parte, los comentarios del docente destacan como el elemento que más ayuda a aprender con un 60,5% de las respuestas (Tabla 4), seguido por el 30,2% de la nota numérica. Ese dato refleja la importancia del feedback individualizado en el proceso de evaluación, considerado un elemento esencial en el aprendizaje formativo (Sanmartí, 2007). Por otro lado, las opciones que implican la participación de los compañeros representan solo un 9,3% del total, lo que puede indicar la limitada experiencia del alumnado en estos tipos de evaluación o la falta de confianza en estos (Strijbos & Wichmann, 2018).

En cuanto a la utilidad percibida de la autoevaluación y coevaluación (Tabla 5), se observa como la autoevaluación es considerada útil siempre en un 60,5% de los casos, frente al 30,2% de la coevaluación. Esta diferencia puede mostrar una mayor confianza del alumnado en su capacidad de evaluación frente a la de sus compañeros, algo que queda evidenciado en estudios que destacan que la autoevaluación bien implementada puede fomentar la autorregulación y metacognición (Andrade & Du, 2007). Sin embargo, los datos muestran que la coevaluación no es rechazada ya que un 55,8% aceptan su utilización en algunos casos.

La siguiente variable estudiada es la utilización de estas modalidades de evaluación en la nota final (Tabla 6), donde un 55,8% están a favor, tanto en autoevaluación como en coevaluación. No obstante, gran parte del alumnado solo las acepta si coinciden con la opinión del docente. Esto refleja que la apertura a nuevas formas de evaluación requieren todavía de validación externa, ya que la confianza del alumnado en su propio juicio o el de sus iguales se encuentra en desarrollo y requiere del acompañamiento docente para ganar legitimidad (Tirado González, 2009).

Adentrándonos en el análisis de asociación entre variables (Tabla 7), se observa que el grupo 4°C prefiere que el profesor sea el único agente evaluador, mientras que los grupos 4ºA+DIV y 4ºB muestran una distribución más variada en sus respuestas, es algo destacable ya que es el grupo que participó activamente en la coevaluación, demostrando por tanto que no por ello han interiorizado los beneficios de esta práctica, tal y como se sugiere en el estudio de Cervantes y Medina (2020). Al mismo tiempo, todos los grupos muestran una preferencia por la evaluación continua a lo largo de la Unidad Didáctica (Tabla 8), lo que podría ser una tendencia hacia modelos que no se centran únicamente en un producto final, sino que valoran el proceso de aprendizaje completo (Pérez Pino

et al., 2017).

En cuanto a la inclusión de estas evaluaciones en la nota final (Tabla 10), todos los grupos coinciden en la aceptación de la autoevaluación y coevaluación como parte de la calificación final de la asignatura. Esto sugiere que a pesar de las diferentes formas de trabajo llevadas a cabo en cada grupo, todos ellos están abiertos hacia estas metodologías (Cervantes & Medina, 2020).

Por su parte, en la opinión general sobre la autoevaluación (Tabla 11), las medias obtenidas para las variables descritas como por ejemplo la capacidad de evaluar sin ayuda externa o la conciencia de progreso, son bastante altas, entre  $\bar{x}=3,74$  y  $\bar{x}=4,14$ , situándose por tanto en el valor de “Casi siempre”. Aunque no se alcanzan niveles de significación, desde un punto de vista didáctico, puede resultar interesante incluso para tener en cuenta para futuros estudios como se observa que el grupo que practicó la autoevaluación (4ºB) presenta valores ligeramente superiores, lo que apoya los beneficios de esta práctica cuando se integra de forma sistemática y estructurada (Andrade & Du, 2007).

De la misma forma, para la opinión general sobre la coevaluación (Tabla 12), las medias de todas las variables excepto una, son positivas, destaca únicamente la variable de dificultad para ser objetivo con una media de  $\bar{x}=3,02$  que equivale a la categoría “A veces” y que por tanto sugiere que esta práctica presenta este tipo de limitaciones a la hora de ser puesta en práctica. La influencia de factores afectivos, como la amistad entre compañeros, puede dificultar una valoración objetiva, tal y como han documentado estudios previos sobre evaluación entre iguales (Strijbos & Wichmann, 2018).

Para finalizar con el análisis cuantitativo, ponemos el foco en el análisis de correlaciones. En primer lugar, la Tabla 14 muestra como existe una correlación negativa, aunque no significativa entre el agente evaluador y tanto la utilidad de la autoevaluación ( $p = 0,246$ ) como con la utilidad de la coevaluación ( $p = 0,604$ ). Esto indica que no existe una asociación entre quién se prefiere que evalúe (profesor, alumno o compañeros) y la utilidad que se le atribuye a estos modelos de evaluación. Sin embargo, la dirección negativa de los coeficientes sugiere que quienes prefieren modelos más tradicionales tienden a valorar menos la evaluación compartida, lo que concuerda con estudios que identifican resistencias en alumnado habituado a la heteroevaluación (Tirado González, 2009).

En la correlación entre el elemento que más ayuda a aprender y la utilidad de la autoevaluación y coevaluación (Tabla 15), se observa una baja correlación negativa pero no significativa para la autoevaluación ( $p=0,643$ ) y una leve correlación positiva, aunque tampoco significativa para la coevaluación ( $p=0,121$ ) por lo que no puede afirmarse una relación entre estas variables. Aunque no son significativas, estas correlaciones reflejan que los estudiantes que valoran más los comentarios del profesor y tienden también a mostrar una actitud más receptiva hacia la coevaluación, lo cual sugiere que el feedback es percibido como un recurso útil, en la evaluación (Pérez Pueyo et al., 2008).

Los valores que nos aporta la Tabla 16 muestran la correlación entre la utilización de la autoevaluación para la nota y su utilidad, obteniendo una correlación positiva, aunque no significativa ( $p=0,061$ ). Esto indica una tendencia de que los estudiantes que consideran útil la autoevaluación también aceptan su inclusión en la nota final. Aunque no es significativo, apoya la idea de que cuando el alumnado reconoce la utilidad de una herramienta, aumenta su disposición para integrarla en procesos más formales como en este caso la calificación (Andrade & Du, 2007).

La correlación entre la utilización de la coevaluación para la nota y su utilidad (Tabla 17) presenta una correlación positiva débil y no significativa ( $p=0,193$ ), siguiendo la misma línea de lo expuesto en el caso de la autoevaluación.

La única correlación estadísticamente significativa con  $p=0,022$  es la que se muestra en la Tabla 18 y que relaciona la percepción de ayuda al progreso de la autoevaluación y su utilidad. En este caso, se trata de una correlación negativa que considera por tanto que los alumnos a los que la autoevaluación les ayuda a ser conscientes de su progreso, la perciben como menos útil. Aunque pueda parecer una contradicción, esto puede deberse a una elevada autoexigencia, donde la conciencia de los errores y las debilidades puede generar una percepción más crítica sobre uno mismo (Gutiérrez-García et al., 2011)

Como último análisis correlacional, la Tabla 19 muestra la correlación entre la percepción de aprendizaje en la coevaluación y su utilidad, siendo esta negativa y no significativa ( $p=0,256$ ) que indica que los alumnos que consideran que han aprendido evaluando consideran menos útil la coevaluación. Esto, aparte de evidenciar la falta de formación para la coevaluación, puede indicar que aunque piensen que aprenden evaluando, consideran este tipo de evaluación no es lo suficientemente riguroso como para ser utilizado en una evaluación formal (Topping, 2009).

Para finalizar, el análisis cualitativo se ha realizado siguiendo el enfoque metodológico de Guillén Correas (2014). En primer lugar, centrándonos en la autoevaluación (P10), una de las categorías más destacadas fue la objetividad con respuestas como “ser más objetivo” (122), “ser más objetiva conmigo misma y aceptar las críticas” (111) o “ser objetivo y rectificar mis errores” (234). Junto a esta categoría encontramos las categorías de aceptación de errores y búsqueda de apoyo externo como respuestas más recurrentes, algo que concuerda con lo estudiado en diferentes artículos, que dicen que la autoevaluación contribuye al desarrollo de la autorregulación cuando se asocia a criterios claros y explícitos (Panadero et al., 2016). También son recurrentes aunque en menor medida las categorías de falta de experiencia y uso de herramientas como por ejemplo la grabación de vídeos, lo cual destaca la importancia de una reflexión basada en procesos de evaluación personal (Boud et al., 2013).

No obstante, cabe destacar que un número significativo de respuestas fueron del tipo “no sé” (114) o “nada” (320), lo que puede interpretarse como una falta de reflexión crítica o un escaso desarrollo de la competencia evaluadora (Carrillo, 2016).

Por otro lado, en referencia a la coevaluación (P15), de la misma forma que en la autoevaluación, la categoría que más se repite es la de la objetividad, lo que demuestra que en contexto donde predominan las relaciones personales, la evaluación entre iguales supone desafíos para lo que se requiere entrenamiento específico y experiencia previa (Falchikov, 2003). Otras categorías muy repetidas son la concentración (“fijarme más” (203), “prestar mucha atención a cómo juegan” (243)), reconocer el esfuerzo ayudar al compañero.

Finalmente, al igual que en la autoevaluación, aparecen respuestas como “no sé” (201) o “creo que ya lo hago bien” (238), que podrían evidenciar una falta de reflexión crítica o un escaso desarrollo de la capacidad evaluadora.

## **7. CONCLUSIONES**

En este apartado, vamos a exponer las principales conclusiones a las que hemos llegado tras la realización del presente estudio.

1. Este trabajo ha permitido corroborar que, pese a que la heteroevaluación sigue siendo la más reconocida por el alumnado como método más justo para la evaluación, existe una apertura hacia modelos de evaluación más participativos como la autoevaluación y la coevaluación. Esta tendencia, aunque todavía no es clara, podría evidenciar una transición hacia enfoques más formativos y centrados en el proceso de aprendizaje.

2. La mayoría del alumnado prefiere una evaluación continua a lo largo de toda la Unidad Didáctica, lo que se alinea con las características de la evaluación formativa, destacando su importancia como herramienta de mejora continua y no únicamente a los resultados finales.

3. En cuanto a la percepción de utilidad, la autoevaluación es mejor valorada que la coevaluación, lo que puede deberse a una mayor confianza del alumnado en su propio juicio sobre el de sus compañeros. No obstante, ambas son ampliamente aceptadas para tenerlas en cuenta en la calificación final, especialmente cuando sus resultados coinciden con los del docente.

4. A pesar de que no se encontraron diferencias significativas entre grupos respecto a la valoración general de los diferentes tipos de evaluación estudiados, se observa una ligera tendencia de los grupos que han experimentado autoevaluación y coevaluación a estar más a favor de estas prácticas. Esto sugiere que el uso continuado de estas estrategias puede favorecer a mejorar la percepción en ellas y es un tema a tener en cuenta para futuras investigaciones.

5. El análisis cualitativo pone en evidencia que gran parte del alumnado identifica la objetividad, la capacidad crítica y el reconocimiento del error como elementos clave para mejorar su capacidad de autoevaluación y coevaluación. También destaca la falta de experiencia en este tipo de prácticas.

## **8. LIMITACIONES DEL ESTUDIO**

La principal limitación de este Trabajo de Fin de Máster es el tamaño de la muestra, que se compone únicamente de 43 alumnos y alumnas de 4º de la ESO, lo que representa un porcentaje muy bajo del alumnado de estas características que podría ser estudiado. Esta baja participación puede generar que los datos obtenidos, pese a poder ser representativos del total de alumnos y alumnas de 4º de la ESO, se deben interpretar con cautela las conclusiones obtenidas, sobre todo si se busca extrapolar el estudio a un ámbito más grande, para lo que sería necesario llevar a cabo un estudio más exhaustivo que recogiera un número suficientemente representativo del alumnado de estas características.

De la misma forma, el estudio se desarrolló en un único centro educativo y con un único nivel académico, lo cual limita la diversidad de participantes y puede influir en la homogeneidad de las respuestas, para ello, de igual manera que en el caso del tamaño de la muestra, sería necesario llevar a cabo otro estudio en el que se incluyeran otros centros e incluso otros niveles educativos.

Por otro lado, algunas preguntas abiertas del cuestionario mostraron una falta de reflexión por parte del alumnado, lo que limita la riqueza del análisis cualitativo. Esta situación podría deberse a una escasa familiarización del alumnado con este tipo de prácticas.

Por último, es importante tener en cuenta que no se llevó a cabo ninguna medición previa a la intervención, por lo que no es posible afirmar si las percepciones y actitudes del alumnado hacia estas prácticas evaluativas se modifican tras su aplicación práctica en el proceso de intervención.

## 9. REFERENCIAS BIBLIOGRÁFICAS

- Alstot, A. E. (2018). Accuracy of a Peer Process Assessment Performed by Elementary Physical Education Students. *The Physical Educator*, 75(5), 8293. <https://doi.org/10.18666/TPE-2018-V75-I5-8293>
- Álvarez Méndez, J. M. (2002). Evaluar para conocer, examinar para excluir. *Evaluar Para Conocer, Examinar Para Excluir*.
- Amaya Cedrón, L. A. (2017). Prueba Chi-Cuadrado en la estadística no paramétrica. *Ciencias*, 1(1), 13–17.
- Andrade, H., & Du, Y. (2007). Student responses to criteria-referenced self-assessment. *Assessment & Evaluation in Higher Education*, 32(2), 159–181.
- Araujo, P. J. (2014). Manual introductorio al uso del programa PSPP para el análisis de datos. *Biblioteca Las Casas*. Recuperado de: <Http://Www.Indexf.Com/Lascasas/Documentos/Lc0773.Php>.
- Backman, E., Quennerstedt, M., Tolgfors, B., & Nyberg, G. (2024). Peer assessment in physical education teacher education – a complex process making social and physical capital visible. *Curriculum Studies in Health and Physical Education*, 15(3), 274–288. <https://doi.org/10.1080/25742981.2023.2256327>
- Bernardino, L., López-Pastor, V. M., & Monjas, R. (2006). Evaluación formativa y compartida en Educación Física. *Ágora Para La EF y El Deporte*, 3, 75–99.
- Boud, D., Lawson, R., & Thompson, D. G. (2013). Does student engagement in self-assessment calibrate their judgement over time? *Assessment & Evaluation in Higher Education*, 38(8), 941–956.
- Cañadas, L., & Santos-Pastor, M. L. (2021). La evaluación formativa desde la perspectiva de docentes noveles en las clases de educación física en primaria y secundaria. *Revista Electrónica Educare*, 25(3), 452–471.
- Carrillo, L. S. G. (2016). La autoevaluación de los estudiantes un proceso por resignificar y reconstruir en la educación física escolar. *Revista Paideia Surcolombiana*, 21, 27–42.
- Cervantes, A., & Medina, J. (2020). Procesos de autoevaluación y coevaluación en Educación Física: una revisión sistemática. *Revista Internacional de Educación y Aprendizaje*, 8(2), 123–

140. <https://revistas.uam.es/riee/article/view/15195>

- Cervantes, M. C., & Medina, J. L. (2020). Dificultades en la implementación de la coevaluación en contextos escolares. *Revista de Investigación Educativa*, 38(1), 129–145.
- Estrada Marcén, N., & Romero Martín, M. . R. (2024). Evaluación formativa en Expresión Corporal: definición de Indicadores de Evaluación a partir de las Reglas de Acción. In S. Congreso Internacional de Educación Física y Ciencias 2024 (Ed.), *Actas del Congreso Internacional de Educación Física y Ciencias*. Universidad de Valladolid / Universidad de Zaragoza.
- Falchikov, N. (2003). *Learning together: Peer tutoring in higher education*. Routledge.
- Fuentes Nieto, T. (2019). Experiencia de evaluación formativa en Educación física en un centro rural en Secundaria. *Revista Infancia, Educación y Aprendizaje*, 5(2), 178–186.
- Galicia Alarcón, L. A., Balderrama Trápaga, J. A., & Edel Navarro, R. (2017). Validez de contenido por juicio de expertos: propuesta de una herramienta virtual. *Apertura (Guadalajara, Jal.)*, 9(2), 42–53.
- Guillén Correas, R. (2014). *Metodología cualitativa aplicada a las ciencias del deporte*. Prensas de la Universidad de Zaragoza.
- Gutiérrez-García, C., Pérez-Pueyo, Á., & López-Pastor, V. M. (2011). Evaluación formativa y autoestima en Educación Física: un estudio de caso. *Revista Electrónica Interuniversitaria de Formación Del Profesorado*, 14(4), 139–147.
- Heras Bernardino, C., Arribas Martín, M. I., Marugán Manrique, J. C., Subtil, P., García, L. M., Pastor, V. M. L., Aguado, R. M., García, J. G., Pastor, E. M. L., & Pinela, J. F. M. (2006). La evaluación en educación física. Revisión de modelos tradicionales y planteamiento de una alternativa. La evaluación formativa y compartida. *RETOS. Nuevas Tendencias En Educación Física, Deporte y Recreación*, 10, 31–41.
- Jiménez Barret, I. (2017). Una experiencia de evaluación formativa en educación física en educación secundaria. *Revista Infancia, Educación y Aprendizaje*, 3(2), 107–111.
- 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)*.

- Mamani Ramos, Á. A. (2017). La evaluación formativa participativa y su impacto en la predisposición por aprender y aprendizaje del área de Educación Física en estudiantes del segundo grado de Educación Secundaria. *Sportis*, 3(1), 206–220.
- McKight, P. E., & Najab, J. (2010). Kruskal-wallis test. *The Corsini Encyclopedia of Psychology*, 1.
- Mendivelso, F. (2021). Prueba no paramétrica de correlación de Spearman. *Revista Médica Sanitas*, 24(1).
- Miras, F., & Granda, M. (2018). Resultados de un caso práctico de autoevaluación en Educación Física. *Retos: Nuevas Tendencias En Educación Física, Deporte y Recreación*, 34, 22–27. <https://www.redalyc.org/pdf/5516/551656959004.pdf>
- Panadero, E., Brown, G., & Strijbos, J.-W. (2016). The Future of Student Self-Assessment: a Review of Known Unknowns and Potential Directions. *Educational Psychology Review*, 28, 803–830. <https://doi.org/10.1007/s10648-015-9350-2>
- Pérez Pino, M, González-Gijón, G., & López-Pastor, V. M. (2017). La evaluación formativa en Educación Física: percepciones del alumnado. *Retos*, 32, 217–223.
- Pérez Pino, Michel, Enrique Clavero, J. O., Carbó Ayala, J. E., & González Falcón, M. (2017). La evaluación formativa en el proceso enseñanza aprendizaje. *Edumecentro*, 9(3), 263–283.
- Pérez Pueyo, Á., Heras Bernardino, C., & Herrán Álvarez, I. (2008). Evaluación formativa en la educación secundaria obligatoria. Su aplicación a una unidad didáctica de deportes colectivos en el marco del estilo actitudinal. *Revista Española de Educación Física y Deportes*, 383, 45.
- Razali, N. M., & Wah, Y. B. (2011). Power comparisons of shapiro-wilk, kolmogorov-smirnov, lilliefors and anderson-darling tests. *Journal of Statistical Modeling and Analytics*, 2(1), 21–33.
- Rendón-Macías, M. E., Villasís-Keever, M. Á., & Miranda-Novales, M. G. (2016). Estadística descriptiva. *Revista Alergia México*, 63(4), 397–407.
- Sánchez-Martínez, A., Moreno, J. L., & Ion, G. (2019). Diseño y Validación de un Cuestionario de Percepción del Aprendizaje a través del Feedback entre Iguales en Educación Superior. *Revista Iberoamericana de Diagnóstico y Evaluación-e Avaliação Psicológica*, 4(53).
- Sanmartí, N. (10 C.E.). Ideas clave. *Evaluar Para Aprender*. Madrid: Ed. Graó.

- Schettini, P., & Cortazzo, I. (2016). Técnicas y estrategias en la investigación cualitativa. *Series: Libros de Cátedra*.
- Strijbos, J.-W., & Wichmann, A. (2018). Promoting learning by leveraging the collaborative nature of formative peer assessment with instructional scaffolds. *European Journal of Psychology of Education*, 33(1), 1–9.
- Tirado González, J. (2009). La autoevaluación en Educación Física. *Lecturas: Educación Física y Deportes*, 13(130). <https://dialnet.unirioja.es/descarga/articulo/3175446.pdf>
- Topping, K. J. (2009). Peer assessment. *Theory Into Practice*, 48(1), 20–27.  
<https://doi.org/10.1080/00405840802577569>
- Vázquez Avilés, N. (2019). La evaluación educativa como estrategia didáctica. *Atlante Cuadernos de Educación y Desarrollo*, febrero.

## ANEXO I: TABLA ÍTEMS DEL CUESTIONARIO

Puede consultar el cuestionario completo a través del siguiente enlace:

<https://forms.gle/QiCnv6N754RpoFwV9>

<b>Nº Pregunta</b>	<b>Enunciado</b>
1	Indique su edad
2	Indique su sexo
3	¿Cómo prefieres que te evalúen?
4	¿Cuándo prefieres que te evalúen?
5	¿Qué te ayuda más a aprender?
6	¿Te parece útil evaluar tu propio trabajo?
7	¿Qué parte te resulta difícil al autoevaluarte en una actividad?
8	¿Te parece justo que el profesor/a tenga en cuenta tu autoevaluación para la nota final?
9	Seleccione la opción que mejor se adapte a tu opinión sobre la autoevaluación
10	¿Qué harías para mejorar tu forma de evaluarte?
11	¿Te parece útil evaluar el trabajo de otros?
12	¿Cómo prefieres dar tu opinión sobre lo que hace un compañero?
13	¿Te parece justo que el profesor/a tenga en cuenta la coevaluación para la nota final?
14	Seleccione la opción que mejor se adapte a tu opinión sobre la coevaluación
15	¿Qué harías para mejorar tu forma de evaluar a otros en clase?

## ANEXO II. TABLAS ANÁLISIS CUALITATIVO

P10. Cómo mejorar la autoevaluación	
Categoría	Respuesta
Objetividad	Ser más objetivo, Ser objetivo, Ser más objetiva conmigo misma y aceptar las críticas, Ser más objetiva, Ser objetivo y rectificar mis errores, ser mas objetivo
Aceptación del error	Darme cuentas de mis errores antes, aceptar errores, Fijarme más en mis errores, aceptar los errores, no ser tan autoexigente a veces
Búsqueda de apoyo externo	Preguntar a mis compañeros, Pedir consejos a el profesor o compañeros, Consultar con mis compañeros, Añadir más consejos del profesor, Compararme con otras personas y ejemplos, Compararme con alguien que se le dé bien y así poder fijarme y mejorar más
Falta de experiencia	Ser más consciente de lo que hago mientras juego, Practicar más, Ser más consciente de lo que tengo que hacer y hago mientras juego, Hacerlo más, evaluarme más
Uso de herramientas	Grabarme, Grabrne
Indefinición	Nose, no se, Ns, Pues no lo sé la verdad, no se, Nada, No se, nada, No se, No sé, No lo sé e ver mis errores, .
Otros	Hablarlo, Seguir aprendiendo, Preguntar más o piñones, Tener en cuenta que no veo muy bien el volante

P15. Cómo mejorar la coevaluación	
Categoría	Respuesta
Objetividad	Ser más objetivo, Ser más objetiva y no tan subjetiva..., Ser más objetiva sin tener en cuenta las amistades, Ser más objetivo con los que son mis amigos, Ser una persona más objetiva y pensar bien lo que voy a poner, Ser objetiva
Concentración	fijarme más, Mirar como lo hacen, Estar atenta y evaluar de forma objetiva..., Estar más atento a los demás, prestar mucha atención a cómo juegan
Reconocimiento del esfuerzo	Ser consciente del esfuerzo de cada persona..., Valorar el esfuerzo del otro compañero
Ayuda al compañero	Ayudarles, Pensar en mis fallos y poder ayudar a los demás, Pidiendo consejos también..., darles consejos para que lo puedan hacer lo mejor posible
Uso de instrumentos estructurados	usar una rúbrica
Indefinición	Nose, No lo sé, No lo sé, No lo se, No se, Nada, nada, No sé nada creo que así está bien, No se, nada, creo que ya lo hago bien, ,, .
Otros	Que el profesor diga lo que fallas, Saber más sobre el tema que estamos dando, E pensar que estamos aprendiendo y que no somos profesionales, Consultar con mis compañeros