



Impacto de un debate sociocientífico en las habilidades argumentativas y en la toma de decisiones del profesorado de secundaria en formación inicial

Impact of a Socio-Scientific Debate on the Argumentative and Decision-Making Skills of Preservice Secondary Teachers

Pilar Bernal-Herrera, María José Cano-Iglesias, Antonio Joaquín Franco-Mariscal, Ángel Blanco-López
Didáctica de las Ciencias Experimentales. Universidad de Málaga. Málaga. España
pilar_9@uma.es; mjcano@uma.es, anjoa@uma.es, ablancol@uma.es

RESUMEN • Esta investigación explora el impacto de un debate sobre la elección de un producto de la vida diaria en las habilidades de argumentación y en la toma de decisiones, dimensiones importantes del pensamiento crítico, en 45 estudiantes de especialidades científicas del Máster en Profesorado de Educación Secundaria. Se presenta el caso sobre el material (acero o aluminio) que prefieren al comprar una bicicleta empleando como instrumentos un pretest y un postest sobre decisiones y argumentos de oyentes, así como sobre los informes de preparación y grabación en audio del debate. Se analizan las decisiones de los oyentes antes y después y sus posibles cambios, además de los argumentos dados, empleando una rúbrica que incluye las pruebas, la justificación y la conclusión. Tras el debate se producen progresos, estadísticamente significativos en los oyentes, con respecto al uso de argumentos con un mayor número de pruebas, una tipología más amplia de ellas y una mejor calidad de las justificaciones y conclusiones empleadas. Las mejoras son más relevantes en estudiantes de grados no afines que afines al estudio de la materia.

PALABRAS CLAVE: Argumentación; Toma de decisiones; Pensamiento crítico; Debates; Formación inicial de profesorado.

ABSTRACT • This research explores the impact that a debate about the choice of an everyday product has on argumentation and decision-making skills, which are important dimensions of critical thinking, among 45 students specializing in sciences in the master's degree in Secondary Education. This case involves a choice between materials (steel or aluminum) when purchasing a bicycle, using as instruments a pre- and post-test to assess listeners' decisions and arguments, as well as the preparation reports and audio recordings of the debate. The listeners' decisions before and after the debate, together with their potential changes, and the arguments presented are analyzed using a rubric that includes the evidence, the justification, and the conclusion. Significant progress is observed in listeners after the debate, with regards to the use of arguments supported by a greater number of evidence, a wider range of evidence types, and enhancements in the quality of justifications and conclusions. This progress is more significant among students with unrelated degrees to the subject of study.

KEYWORDS: Argumentation; Decision-making; Critical thinking; Debates; Initial teacher training.

Recepción: diciembre 2022 • Aceptación: septiembre 2023 • Publicación: noviembre 2023

FUNDAMENTACIÓN TEÓRICA

La investigación que se presenta en este artículo considera que el desarrollo de la toma de decisiones y de la argumentación en ciencias, en cuanto que aspectos destacados del pensamiento crítico, son objetivos importantes de la educación científica que se pueden abordar en el contexto de problemas de la vida diaria, particularmente en el de la elección de productos, y se entiende que el debate puede ser una estrategia muy adecuada para su desarrollo. A continuación, se presentan y desarrollan estas ideas.

Pensamiento crítico en la enseñanza de las ciencias

Hoy día existe un acuerdo unánime sobre la importancia que el pensamiento crítico tiene para los ciudadanos y sobre su consideración como una de las grandes finalidades de la educación científica (Osborne, 2014). Desafortunadamente, durante muchos años la educación científica ha olvidado el desarrollo de habilidades de pensamiento crítico (Osborne, 2001), a pesar de que la sociedad actual exige que la enseñanza de las ciencias juegue un papel activo en su desarrollo en el contexto de problemas relacionados con aspectos energéticos, ambientales, alimentación o salud. Por ello, el pensamiento crítico constituye un reto para la educación científica (Duschl et al., 2007; Jenkins, 2011) y, cada vez más, la literatura especializada presta atención a este tema (Torres y Solbes, 2016; Jiménez y Otero, 2019; Bargiela et al., 2022).

Osborne (2014) apunta que crítica y cuestionamiento son esenciales para la práctica de la ciencia y que sin ellos no se podría construir conocimiento científico. Jiménez-Aleixandre (2010) entiende el pensamiento crítico como la capacidad de desarrollar una opinión independiente, adquiriendo la facultad de reflexionar sobre la sociedad y participar en ella, e incluye la búsqueda y el uso de pruebas como componentes importantes de la argumentación científica. Al respecto, Solbes y Torres (2012) definen el pensamiento crítico como una necesidad para que las personas estructuren una manera de pensar propia y sean capaces de tomar posiciones frente a las situaciones sociales que viven y poder tener un papel activo en las decisiones culturales y científicas.

Aunque el pensamiento crítico puede desarrollarse desde la enseñanza de las ciencias mediante el tratamiento de cualquier tipo de problema (Bailin, 2002), existe acuerdo en la relevancia de los problemas de la vida diaria en los que la ciencia y la tecnología juegan un papel importante, especialmente sobre los que existen controversias (Torres y Solbes, 2016) y en los que la ciudadanía necesita tomar decisiones que les afectan personal y socialmente. Partiendo de la caracterización que Solbes y Torres (2012) hacen de las competencias que se requieren para el desarrollo del pensamiento crítico en el ámbito de problemas sociocientíficos, Blanco et al. (2017) identificaron un conjunto de dimensiones para su desarrollo en la enseñanza de las ciencias. Entre ellas, la toma de decisiones y la argumentación destacan como habilidades fundamentales.

Toma de decisiones y argumentación sobre problemas de la vida diaria

Desde la educación científica se considera fundamental formar al alumnado en la toma de decisiones sobre situaciones y problemas de la vida diaria cuyas consecuencias pueden afectar a la ciudadanía en distintos ámbitos. De ahí la importancia de disponer de estudiantes alfabetizados científica y tecnológicamente capaces de ejercer plenamente sus derechos y participar en la toma de decisiones en las sociedades democráticas actuales. Por tanto, esta formación tiene que ayudarles a adoptar posturas responsables, argumentadas y basadas en conocimientos científicos y tecnológicos (Sadler y Zeidler, 2005; Yacoubian, 2018).

En la toma de decisiones en la vida diaria intervienen factores relacionados con las preferencias y creencias personales y vinculados al entorno (Muñetón et al., 2017). Es aquí donde el pensamiento crítico es relevante (es decir, analizar, reflexionar, seleccionar información, etc.) para tomar una decisión lo más adecuada posible al problema planteado.

Las situaciones de la vida real constituyen contextos útiles para el desarrollo de aprendizajes (Lupión et al., 2017), ya que al tratarse de problemas reales aumenta la probabilidad de poner en práctica valores y actitudes que normalmente se usan en la toma de decisiones de nuestra vida diaria (Prieto et al., 2012). Acevedo (2006) afirma que por medio de controversias sobre temas tecnocientíficos de interés social se puede observar claramente la relación entre educación en valores y enseñanza de las ciencias. Según López (2012), el uso de problemas sociocientíficos ofrece beneficios al alumnado, como un mayor dominio y retención de la materia, una mayor habilidad para el posicionamiento del estudiante o una mejor toma de decisiones.

Diferentes estudios revelan que las principales dificultades que encuentra el alumnado para la toma de decisiones sobre problemas de la vida diaria se asocian con la falta de conocimientos científicos para argumentar adecuadamente (Henderson et al., 2018; Rodríguez-Mora et al., 2022) o con el uso frecuente de creencias o valores al tomar decisiones (von Winterfeldt, 2013; Moreno-Fontiveros et al., 2022). Este hecho es preocupante en la didáctica de las ciencias, que intenta enseñar a tomar decisiones argumentadas basadas en pruebas científicas (Bravo y Jiménez-Aleixandre, 2018), en actitudes y valores sociales (Siribunnam et al., 2014) y en el uso de justificaciones basadas en conocimientos científicos.

Así pues, la argumentación constituye un aspecto importante para la toma de decisiones (Bravo y Jiménez-Aleixandre, 2018) y requiere de la construcción y la crítica de argumentos (Osborne et al., 2016). El modelo de Toulmin (2003) o Toulmin's Argument Pattern es uno de los más utilizados para analizar argumentos. Según Erduran et al. (2004, p. 57), además de una conclusión, el argumento debe incluir «pruebas que apoyen esa afirmación; justificaciones que proporcionen un vínculo entre las pruebas y la afirmación; respaldos que refuercen las justificaciones; y, por último, refutaciones que señalen las circunstancias en las que la afirmación no sería cierta». La literatura muestra que el alumnado encuentra dificultades al identificar y usar pruebas en los argumentos, y sus progresos en este sentido pueden considerarse un indicador de la calidad de los argumentos empleados (Bravo y Jiménez-Aleixandre, 2018).

Argumentar sobre problemas de la vida diaria implica procesos de razonamiento informal. Este razonamiento se aplica fuera de contextos formales, en los que los problemas no están bien definidos, las premisas pueden no estar explícitamente enunciadas y las conclusiones pueden no estar delimitadas (Wu, 2013; Cruz-Lorite et al., 2023). Diversos autores indican diferentes tipos de razonamiento informal según los dominios a los que se orientan los argumentos (p. e., científico-tecnológico, económico, social, etc.) (Christenson et al., 2012; Ozturk y Yilmaz-Tuzun, 2017). El uso y prevalencia de estos tipos de razonamiento pueden estar influenciados por el problema o los conocimientos previos.

La elección de productos como contexto para la toma de decisiones y la argumentación

La elección de productos es un contexto interesante para tomar decisiones y argumentar, pues en ellas influyen distintos factores, entre los que destaca la publicidad (Grosick et al., 2013; Girón y Lupión, 2022). Por ejemplo, los jóvenes eligen determinados productos alimenticios en función de su publicidad, por lo que esta impacta en estilos de vida y comportamientos (Villani, 2001).

Rivera et al. (2009) definieron el comportamiento del consumidor como el «proceso de decisión y la actividad física que los individuos realizan cuando buscan, evalúan, adquieren y usan o consumen bienes, servicios o ideas para satisfacer sus necesidades» (p. 36). Así, el consumidor debe combinar los

estímulos recibidos (producto, precio, promoción, marca) con sus intereses individuales (modelo, uso) para decidir satisfactoriamente (Blanco y Forero, 2017).

Un contexto de la vida diaria relevante para los estudiantes puede ser la elección de un coche (Moreno-Fontiveros et al., 2022) u otros medios de transportes, como la compra de una bicicleta, que constituye el debate del estudio de caso de este trabajo. Se trata de un contexto relevante para promover, entre otras, la conciencia ambiental (Hadjichambis et al., 2020; Moreno-Fontiveros et al., 2015).

La compra de una bicicleta incluye varias vertientes que el alumnado debería considerar en su decisión en cuanto a una elección adaptada a sus necesidades, teniendo en cuenta el impacto que tiene su compra y uso desde el punto de vista social, ambiental y global, o desarrollando un pensamiento crítico que le permita interpretar la información de los mensajes publicitarios (Moreno-Fontiveros et al., 2022).

Hoy día, con el crecimiento masivo de las ciudades y la mayor concienciación ambiental, muchas personas optan por el uso de la bicicleta como alternativa ecológica al transporte. Otras personas la utilizan para hacer deporte. Independientemente de su uso, la decisión del material de la bicicleta es importante.

El debate como estrategia didáctica para la toma de decisiones y la argumentación

El debate se considera una actividad adecuada para desarrollar pensamiento crítico desde problemas de la vida diaria, al basarse en el razonamiento, la argumentación, la toma de decisiones y la comunicación (Fang et al., 2019; Simonneaux, 2001). El debate requiere que el alumnado evalúe e identifique datos e información, reflexione sobre diferentes opiniones, defienda y razone de forma argumentada su posición y tome decisiones basadas en pruebas de calidad (Wang et al., 2017). Facilita de este modo la exposición desde diferentes perspectivas, convirtiéndose en una buena estrategia no solo porque mejora la adquisición de competencias y aprendizajes, sino también porque favorece la autonomía y la participación, por lo que contribuye a una formación integral del estudiantado, que desarrolla conocimientos y actitudes de manera transversal e interdisciplinar (Rodríguez, 2012).

Partiendo del uso de debates sobre problemas de la vida diaria, en este caso sobre la elección de una bicicleta en función del material con la que está fabricada, la novedad de esta investigación radica en que se centra en el impacto que esta estrategia didáctica puede tener en los oyentes, tanto en los criterios utilizados para adoptar la decisión como en la mejora de los argumentos empleados para justificarla, así como en su formación previa afín o no al caso planteado. La hipótesis de partida es que no solo los participantes, sino también los oyentes en un debate, mejoran sus argumentos con un mayor número, tipología y calidad de pruebas, y emplean justificaciones más elaboradas para apoyar su conclusión.

OBJETIVOS Y PREGUNTAS DE INVESTIGACIÓN

Este trabajo se centra en el desarrollo de la argumentación y la toma de decisiones como dimensiones del pensamiento crítico mediante debates sobre problemas de la vida diaria en profesores de secundaria en formación inicial (en adelante, PFI) mostrando los resultados de un estudio de caso sobre la elección del material de una bicicleta.

Se plantean estas preguntas de investigación:

- P1. ¿Qué decisiones adoptan los PFI sobre el problema tratado antes del debate y cómo las argumentan?
- P2. ¿Qué tipos de pruebas y justificaciones utilizan los PFI que intervienen en el debate?
- P3. ¿Cómo cambian, si lo hacen, las decisiones de los PFI y los argumentos que las apoyan después del debate?

METODOLOGÍA

Contexto y participantes

Esta investigación, centrada en el debate «¿Bicicleta de aluminio o de acero?», forma parte de un programa de debates más amplio realizado con 45 PFI (59,5 % mujeres y 40,5 % hombres) de especialidades de ciencias del Máster en Profesorado de Educación Secundaria de la Universidad de Málaga, dentro de la asignatura Innovación Docente e Iniciación a la Investigación Educativa del curso 2021-2022. Se trata de una muestra de conveniencia (Martín y Salamanca, 2007) seleccionada por incluir a PFI con distinta afinidad al problema tratado según su formación académica previa.

Los protagonistas de cada debate son un presentador y dos debatientes con posturas opuestas, mientras que el resto de PFI actúan como oyentes. Todos los PFI intervinieron en un debate durante el desarrollo de la asignatura. Previamente, recibieron tres horas de formación sobre pensamiento crítico y argumentación, que abordó el modelo de argumentación de Toulmin (2003), la versión simplificada de Jiménez-Aleixandre (2010) y el modelo de Osborne et al. (2016). La formación incluía tareas para identificar pruebas, justificaciones y conclusiones en un argumento, así como la elaboración de estos argumentos.

Para el análisis de datos, los participantes se agruparon en PFI cuya formación previa fue un grado afín, es decir, que incluía conocimientos sobre propiedades de materiales (PFI-A) (grado en Química, Ingeniería Química, Biotecnología, etc.) (44,44 %) y aquellos que no los incluían (PFI-NA) (grado en Biología, Veterinaria, Odontología, etc.) (55,56 %). La debatiente a favor de la bicicleta de aluminio disponía de formación afín (Química), mientras que el presentador y la debatiente a favor de acero eran no afines a la materia (Biología).

Descripción de la actividad

Consistió en un debate de corta duración sobre el material más adecuado para comprar una bicicleta (acero o aluminio). El tema se eligió por varias razones. En primer lugar, porque se trata de un vehículo cotidiano muy usado por el estudiantado, lo que permite contextualizar la ciencia en la vida diaria. Por otro lado, porque permite abordar propiedades de materiales e impacto ambiental, incluidos en el currículo de ciencias de secundaria. Por último, la implicación de los PFI en actividades que fomentan la argumentación y la toma de decisiones tiene interés para la transferencia a la práctica educativa.

Antes del desarrollo del debate, los PFI oyentes respondieron la cuestión: «Vas a comprar una bicicleta, ¿cuál elegirías, de acero o aluminio? Razona con argumentos tu decisión».

Para la preparación del debate, debatientes y presentador debían entregar un informe en el plazo de una semana con los argumentos que se utilizarían, indicándose las fuentes consultadas.

El debate se realizó con esta estructura: exposición del problema por el presentador de una forma imparcial mostrando pros y contras de las dos posturas (tres minutos), un minuto inicial para que cada debatiente defendiese su postura, cinco minutos de discusión y un minuto final como conclusión. Tras el debate, se administró a los oyentes la misma cuestión inicial.

Instrumentos de recogida de datos

La toma de datos se realizó antes, durante y después de la actividad con diferentes instrumentos. Las decisiones y argumentos escritos de los oyentes se recogieron con la cuestión planteada como pretest y postest. Los argumentos escritos y orales empleados por las debatientes y el presentador se recogieron a través de los informes y la grabación en audio del debate.

Análisis de datos

Se realizaron dos estudios: el análisis de la toma de decisiones antes y después del debate y sus posibles cambios, y el análisis de argumentos. Ambos tuvieron una parte cualitativa para definir categorías de respuestas y un tratamiento estadístico posterior cuantitativo basado en análisis de frecuencias y establecimiento de comparaciones.

Todos los argumentos, escritos u orales, se analizaron según el modelo de Toulmin (2003) simplificado, identificando pruebas, justificaciones y conclusión. Estos elementos se analizaron de acuerdo con la rúbrica de la tabla 1, adaptada de un estudio anterior (Cano-Iglesias et al., 2021). Las pruebas se analizaron según número, tipo y calidad, considerándose adicionalmente ideas personales.

Tabla 1.
Rúbrica para análisis de argumentos

<i>CONCLUSIÓN</i>					
0: No alcanza ninguna conclusión		1: Duda al alcanzar una conclusión		2: Proporciona una conclusión adecuada y precisa	
<i>PRUEBAS EN CADA ARGUMENTO</i>					
<i>Número</i>					
0: Ninguna prueba	1: Una prueba	2: Dos pruebas	3: Tres pruebas	4: Cuatro pruebas	5: Cinco pruebas
<i>Tipos</i>					
Pruebas económicas (precio)					
Pruebas físicas (peso, densidad)					
Pruebas químicas (resistencia a la corrosión, durabilidad, contaminación ambiental)					
Pruebas mecánicas (resistencia mecánica, a golpes, fabricación, reparación)					
Otras pruebas (Consulta a expertos, catálogos, hechos históricos)					
0: Ninguna prueba	1: Una prueba	2: Dos pruebas	3: Tres pruebas	4: Cuatro pruebas	5: Cinco pruebas
<i>Calidad</i>					
0: Las pruebas son incorrectas o inexistentes		1: Alguna prueba incorrecta		2: Pruebas correctas pero imprecisas	
3: Todas las pruebas correctas y precisas					
<i>Ideas personales</i>					
0: No emplea			1: Emplea ideas personales		
<i>JUSTIFICACIÓN</i>					
0: Ninguna justificación		1: Hace una justificación que no relaciona las pruebas con la conclusión		2: Proporciona una justificación que relaciona las pruebas con la conclusión	

La tipología de pruebas incluida en la rúbrica se seleccionó por consenso entre los investigadores autores de este estudio, una vez analizados cualitativamente los argumentos de los PFI. Llama la atención el escaso número de pruebas ambientales empleadas (una), motivo por el que se englobó como prueba química.

Un aspecto que destacar en la aplicación de la rúbrica es que los argumentos tienen una conclusión y, para justificarla, los PFI deben emplear una o más pruebas. Por tanto, el número y tipo de pruebas recogido en el análisis puede ser superior al número de PFI, mientras que la calidad de las pruebas, las

ideas personales y la justificación se analizan para cada argumento como conjunto y coincide con el número de participantes.

El debate se dividió en intervenciones para facilitar el análisis de argumentos. Seguidamente, se ilustra un ejemplo de análisis usando el texto en seminegrita para las pruebas y la cursiva para la justificación.

«El aluminio resiste mucho mejor la corrosión ambiental que el acero. *En climas costeros como en el que estamos es un factor muy importante ya que el acero se oxida y entonces su vida útil podría verse muy reducida*». (Debatiente).

De acuerdo con la rúbrica, este argumento expone una prueba química empleando una justificación (nivel 2) que relaciona dicha prueba con la conclusión, que, en este caso, no se expone explícitamente (nivel 0) al corresponder al debate. La calidad de este argumento es correcta y precisa (nivel 3) y no emplea ideas personales (nivel 0).

Análisis estadístico

Para determinar posibles diferencias estadísticamente significativas en los resultados se realizó, en primer lugar, la prueba de normalidad de Shapiro-Wilk, adecuada para determinar si el conjunto de datos obtenidos provenía o no de una distribución normal y para muestras inferiores a 50 datos, de la que se obtuvo $p < 0,05$ para todas las variables, lo que supone aplicar pruebas no paramétricas.

Posteriormente, se estudió si existían diferencias estadísticas entre las dos muestras independientes (PFI-A y PFI-NA) en todas las variables del pretest y del posttest. En las variables nominales (toma de decisión e ideas personales) se realizó la prueba chi cuadrado de Pearson, y en el caso en el que las frecuencias absolutas esperadas fueron iguales o inferiores a cinco en más allá del 20 % de las casillas, se realizó también la prueba exacta de Fisher. Para el resto de las variables, con carácter ordinal, se realizó la prueba U de Mann-Whitney.

Por último, se estudiaron posibles diferencias significativas entre muestras relacionadas (decisiones y argumentos dados por el propio PFI en pretest y posttest) para las distintas variables, utilizándose la prueba de McNemar para variables nominales y la prueba bilateral de Wilcoxon para variables ordinales. Se estudiaron diferencias para la totalidad de PFI, en función de su afinidad.

El tamaño del efecto de las pruebas de Wilcoxon y U de Mann-Whitney se calculó utilizando la ecuación $r = Z/\sqrt{N}$, donde N es el número de PFI durante los dos momentos de la actividad, y Z el valor de la prueba estadística. En cuanto al valor de r calculado en términos absolutos: 0,1 se considera efecto pequeño, 0,3 efecto medio y 0,5 efecto grande (Cohen, 1988). El software empleado fue SPSS 25.0.

RESULTADOS Y DISCUSIÓN

A continuación, se presentan y discuten los resultados obtenidos organizados en dos apartados: el primero relativo a los PFI debatientes y el segundo a los oyentes.

Pruebas y justificaciones utilizadas por los PFI que intervienen en el debate

Previo a la intervención, el presentador y las debatientes cumplieron un informe con los argumentos que iban a emplear durante el debate. A continuación, se recogen algunos ejemplos:

El acero se comporta mejor ante vibraciones, es más resistente y estético porque utiliza cuadros más clásicos, aunque esto es subjetivo. El aluminio es más ligero. (Presentador).

El aluminio es el material más utilizado en la bicicleta [...] disponemos de más modelos de bicicletas de aluminio que de acero con una estética mejor y más ligera. He traído varios catálogos de tiendas de bicicletas [...] la mayoría son de aluminio [...] (debatiente a favor de aluminio).

Entre sus principales ventajas encontramos su precio reducido [...] por lo que nos ahorraremos dinero. Además, posee gran resistencia mecánica y a la fatiga [...]. Al ser un material con una tracción mecánica tan alta tiene gran resistencia a golpes y deformaciones [...] (debatiente a favor de acero).

Respecto a los argumentos utilizados durante el debate, la tabla 2 recoge una visión global en función del rol de los PFI. Como se observa, la debatiente a favor de la bicicleta de aluminio, a pesar de realizar una intervención más, empleó menos argumentos y pruebas por intervención que la debatiente a favor del acero. Por el contrario, la media de justificaciones fue superior para el aluminio.

Tabla 2.
Resultados generales del debate referidos a las debatientes

	<i>A favor de la bicicleta de</i>	
	<i>Aluminio</i>	<i>Acero</i>
Total de intervenciones	7	6
Total de argumentos	17	19
Media de argumentos por intervención	2,42	3,17
Total de pruebas	26	25
Media total de pruebas por intervención	3,71	4,16
Total de justificaciones	28	26
Media de justificaciones	1,65	1,47

La tabla 3 muestra el análisis de pruebas y justificaciones utilizadas en cada parte del debate. Debido a que cada rol defendía una postura (conclusión) prefijada de antemano, las debatientes centraron sus argumentos, en la mayoría de las ocasiones, en pruebas y justificaciones. En todos los casos, estas justificaciones relacionaban pruebas con la conclusión (implícita) de su rol.

Tabla 3.
Análisis de argumentos empleados durante el debate

		<i>Minuto inicial</i>			<i>Discusión</i>		<i>Minuto final</i>		<i>Total de debatientes</i>
		<i>Pres.</i>	<i>Aluminio</i>	<i>Acero</i>	<i>Aluminio</i>	<i>Acero</i>	<i>Aluminio</i>	<i>Acero</i>	
<i>Argumentos</i>									
Argumentos	Total	5	4	2	9	15	4	2	36
	Intervenciones	1	1	1	5	4	1	1	13
<i>Pruebas</i>									
Número	Total	12	8	5	14	18	4	2	51
Tipo	Económicas	1	0	1	3	2	0	0	6
	Físicas	1	3	0	2	2	1	0	8
	Químicas	4	1	1	0	2	1	1	6
	Mecánicas	2	2	2	5	10	0	0	19
	Otras	4	2	1	4	2	2	1	12

		<i>Minuto inicial</i>			<i>Discusión</i>		<i>Minuto final</i>		<i>Total de debatientes</i>
		<i>Pres.</i>	<i>Aluminio</i>	<i>Acero</i>	<i>Aluminio</i>	<i>Acero</i>	<i>Aluminio</i>	<i>Acero</i>	
Calidad	Nivel 0	0	0	0	0	0	0	0	0
	Nivel 1	0	0	0	4	2	0	0	6
	Nivel 2	4	2	1	3	8	2	2	18
	Nivel 3	8	2	1	2	5	2	0	12
	Media por arg.	2,67	2,50	2,50	1,78	2,20	2,50	2,00	2,17
Ideas personales	Nivel 0	4	2	1	5	11	2	1	20
	Nivel 1	1	2	1	4	4	2	1	14
	Media por arg.	0,20	0,50	0,50	0,44	0,27	0,50	0,50	0,39
<i>Justificación</i>									
Justificación	Nivel 0	-	0	0	2	4	1	1	8
	Nivel 1	-	0	0	0	0	0	0	0
	Nivel 2	-	4	2	7	11	3	1	28
	Media por arg.	-	2,00	2,00	1,56	1,47	1,50	1,00	1,56

El presentador expuso un buen número de pruebas centradas en la historia de la bicicleta (otras) y materiales empleados en su fabricación, comparando propiedades (químicas y, en menor medida, mecánicas).

Las debatientes emplearon pruebas mecánicas, como la resistencia mecánica o a la fatiga (aportadas mayoritariamente por el rol de acero) y pruebas físicas, como la ligereza (peso) u otras, como la información de catálogos de varias empresas de venta de bicicletas (principalmente por el rol de aluminio). Toda esta tipología de pruebas muestra el uso de distintos dominios de razonamiento informal (Wu, 2013). También emergieron ideas personales por parte de ambas debatientes para justificar sus argumentos (von Winterfeldt, 2013). Ninguno de los participantes empleó pruebas relacionadas con el impacto ambiental, la sostenibilidad ni la gestión posterior una vez terminada la vida útil del producto, consideradas importantes en este problema (Moreno-Fontiveros et al., 2022).

La calidad media de las pruebas fue similar en ambos casos en el momento inicial (nivel 2-3), mayor en el caso del acero en la discusión y lo contrario en el minuto final. Durante la discusión afloraron algunas pruebas erróneas (nivel 1) relacionadas con aspectos económicos o mecánicos. Así, por ejemplo, en el debate se transmitió de forma no suficientemente clara a qué tipo de materiales se referían al comparar precios, generalizándose la idea de que las bicicletas de aluminio eran más baratas que de acero.

La utilización de ideas personales también fue idéntica en ambos casos, excepto en la discusión en la que la debatiente a favor de la bicicleta de acero presentó un mayor número de estas.

En todas las partes del debate predominaron justificaciones de nivel 2, siendo empleadas por ambos roles por igual (14/28). La calidad de las justificaciones presenta un perfil similar para ambas debatientes, excepto en el minuto final, en el que fue mayor para la debatiente a favor del aluminio. En este caso, parece que la falta de formación previa sobre las propiedades de los materiales de la debatiente a favor de la bicicleta de acero no supuso, en términos generales, un hándicap para mantener un debate equilibrado con respecto al otro debatiente que sí lo tenía. La preparación del debate parece que ayudó a este aspecto, aunque durante la fase de discusión la debatiente con menos formación previa tuvo que hacer también más uso de ideas personales (Muñetón et al., 2017) para contraargumentar y rebatir.

De cara a la posible influencia del debate en los oyentes, en términos globales, los resultados (tablas 2 y 3) indican que el debate permitió mostrar un buen número de argumentos y diferentes tipos de

pruebas a favor de cada bicicleta (Christenson et al., 2012; Ozturk y Yilmaz-Tuzun, 2017), aunque en mayor medida en el caso de la de acero. Con respecto a la calidad de los argumentos y al uso de ideas personales, puede deducirse que el debate estuvo muy equilibrado.

Decisiones y argumentos de los PFI oyentes antes y después del debate

La tabla 4 presenta los resultados de decisiones y argumentos de oyentes (pretest y postest). Hay que indicar que, en primer lugar, los argumentos se analizaron en función de la afinidad de los PFI y de la opción elegida (acero o aluminio). Este estudio mostró una baja frecuencia de PFI que eligieron acero y reveló que solo dos de ellos mantuvieron su decisión. Por esta razón, se ha optado por presentar un análisis estadístico para el total de PFI, independientemente de su elección.

Tabla 4.
Resultados de los argumentos de los PFI oyentes antes y después del debate

		<i>Pretest</i>			<i>Postest</i>		
		<i>PFI-NA</i> (<i>N</i> = 23)	<i>PFI-A</i> (<i>N</i> = 19)	<i>Total</i> (<i>N</i> = 42)	<i>PFI-NA</i> (<i>N</i> = 23)	<i>PFI-A</i> (<i>N</i> = 19)	<i>Total</i> (<i>N</i> = 42)
Decisión adoptada	Aluminio	20	15	35	18	16	34
	Acero	3	4	7	5	3	8
<i>Argumentos dados</i>							
Pruebas							
Número de pruebas	Total	31	42	73	65	51	116
	Media por PFI	1,35	2,21	1,74	2,83	2,68	2,76
Pruebas económicas	Nivel 0	20	15	35	10	10	20
	Nivel 1	3	4	7	13	9	22
	Nivel 2	0	0	0	0	0	0
	Media por PFI	0,13	0,21	0,17	0,57	0,47	0,52
Pruebas físicas	Nivel 0	4	3	7	5	6	11
	Nivel 1	19	16	35	18	13	31
	Nivel 2	0	0	0	0	0	0
	Media por PFI	0,83	0,84	0,83	0,78	0,68	0,74
Pruebas químicas	Nivel 0	18	11	29	12	10	22
	Nivel 1	5	8	13	11	9	20
	Nivel 2	0	0	0	0	0	0
	Media por PFI	0,22	0,42	0,31	0,48	0,47	0,48
Pruebas mecánicas	Nivel 0	19	11	30	11	10	21
	Nivel 1	4	6	10	11	6	17
	Nivel 2	0	2	2	1	3	4
	Media por PFI	0,17	0,53	0,33	0,57	0,63	0,60
Otras pruebas	Nivel 0	23	15	38	14	13	27
	Nivel 1	0	4	4	8	4	12
	Nivel 2	0	0	0	1	2	3
	Media por PFI	0,00	0,21	0,10	0,44	0,42	0,43

		Pretest			Postest		
		PFI-NA (N = 23)	PFI-A (N = 19)	Total (N = 42)	PFI-NA (N = 23)	PFI-A (N = 19)	Total (N = 42)
Calidad	Nivel 0	1	0	1	0	0	0
	Nivel 1	2	4	6	8	3	11
	Nivel 2	8	3	11	7	3	10
	Nivel 3	12	12	24	8	13	21
	Media por PFI	2,34	2,42	2,38	2,00	2,53	2,24
Ideas personales	Nivel 0	17	13	30	11	14	25
	Nivel 1	6	6	12	12	5	17
	Media por PFI	0,26	0,32	0,29	0,52	0,26	0,40
Justificación							
Justificación	Nivel 0	15	11	26	8	3	11
	Nivel 1	4	0	4	0	0	0
	Nivel 2	4	8	12	15	16	31
	Media por PFI	0,52	0,84	0,66	1,30	1,68	1,48
Conclusiones							
Conclusiones	Nivel 0	0	0	0	0	0	0
	Nivel 1	4	1	5	0	0	0
	Nivel 2	19	18	37	23	19	42
	Media por PFI	1,83	1,95	1,88	2,00	2,00	2,00

Antes del debate

Decisiones

La compra de una bicicleta de aluminio fue la opción más elegida por el 83,33 % de los participantes frente al 16,67 % que prefería acero. El porcentaje de PFI-NA que optó por el aluminio (86,95 %) fue ligeramente superior al de PFI-A (78,94 %).

Algunos argumentos dados para justificar su elección fueron:

Al comprar una bicicleta, elegiría una de aluminio porque no compito y me interesa una bicicleta más ligera, que no se oxide rápido con el paso del tiempo (como sí pasa con el acero), etc. (PFI-NA-09).

Elijo la bicicleta de acero. La de aluminio es más ligera y aguanta mejor la corrosión, pero de precio superior. La de acero es más robusta y se adapta mejor a mi poder adquisitivo actual. (PFI-A-20).

Argumentación

Independientemente de la opción elegida, los PFI-A mostraron un mayor número de pruebas (2,21 pruebas/PFI-A) que los no afines (1,35). La prueba mayoritaria de los afines estaba relacionada con propiedades físicas (0,84 pruebas/PFI-A, usadas por 16/19 PFI), seguidas de pruebas mecánicas (0,53 pruebas/PFI-A, usadas por 8/19 PFI) y químicas (0,42 pruebas/PFI-A, 8/19 PFI), el resto de las pruebas fueron minoritarias. Algunos ejemplos fueron: *La bicicleta de aluminio presenta una ventaja notable debido a que es mucho más ligera y de esta manera es de fácil manejo (PFI-A-32); Es más resistente y duradero, comodidad, mejor para soldar (PFI-A-21).*

Los PFI-NA aportaron pruebas físicas (0,83 pruebas/PFI-NA, usadas por 19/23 PFI) como mayoritarias, siendo el resto minoritarias. Un ejemplo fue: *Es más cómoda una bicicleta de aluminio porque su peso es más ligero* (PFI-NA-18).

La calidad de las pruebas se considera aceptable, predominando las correctas (2,34 puntos/PFI-NA; 2,42 puntos/PFI-A). No prevalecieron las ideas personales (0,26 puntos/PFI-NA; 0,32 puntos/PFI-A).

La mayoría de los PFI no ofreció ninguna justificación (nivel 0), independientemente de su afinidad (65,22 %, PFI-NA; 57,89 %, PFI-A). En todos los casos las conclusiones alcanzadas fueron mayoritariamente adecuadas y precisas (nivel 2) (82,61 %, PFI-NA; 94,74 %, PFI-A). La ausencia de justificaciones hace que las decisiones adoptadas no estén debidamente argumentadas (Sadler y Zeidler, 2005; Yacoubian, 2018).

Los resultados revelan que el perfil de respuesta inicial de los dos grupos participantes era similar, independientemente de su formación. La prueba exacta de Fisher no detectó diferencias estadísticamente significativas en la toma de decisiones ($\chi^2 = 0,68$, $p > 0,05$) ni en ideas personales ($\chi^2 = 0,35$, $p > 0,05$), mientras que la prueba U de Mann-Whitney solo detectó diferencias con un tamaño de efecto medio entre los dos grupos de PFI en el uso de otras pruebas ($Z = -2,286$, $p = 0,022$, $r = 0,35$) y número de pruebas ($Z = -2,625$, $p = 0,009$, $r = 0,41$), ambas a favor de PFI-A. Estas diferencias pueden explicarse porque el problema tratado hizo aflorar en los PFI-A algunos conocimientos científicos relacionados con su formación, por lo que lograron superar su falta de conocimientos científicos al argumentar (Henderson et al., 2018).

Después del debate

Decisiones

Una visión global de las decisiones adoptadas (tabla 2) tras el debate muestra resultados prácticamente idénticos a los del pretest. Así, la prueba exacta de Fisher no mostró diferencias significativas en decisiones ($\chi^2 = 0,71$; $p > 0,05$).

No obstante, se observaron cambios de decisión recogidos en el diagrama Sankey (figura 1), donde se aprecia cómo los PFI-NA son más reacios a estos cambios (34,8 % de los casos) que los PFI-A (15,8 %). Por tanto, como parece lógico, el debate influyó más en los PFI-NA, que no poseían conocimientos sobre propiedades de los materiales (Henderson et al., 2018; Rodríguez-Mora et al., 2022).

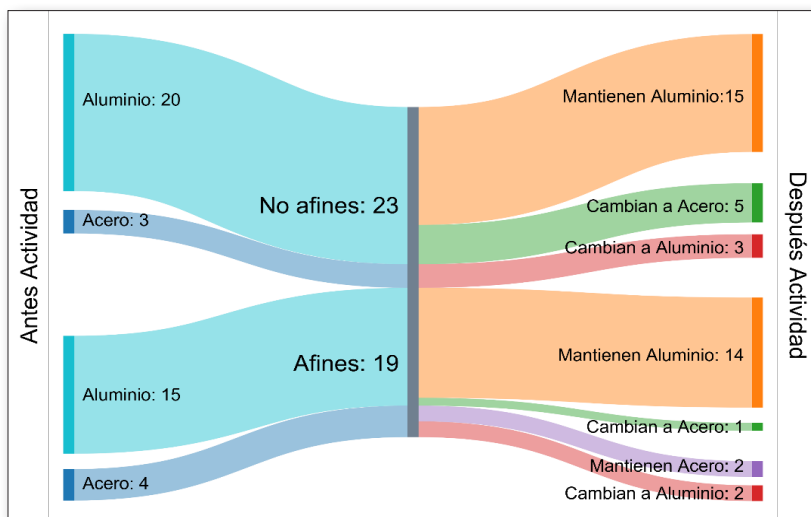


Fig. 1. Diagrama Sankey mostrando cambios de decisión.

Algunos argumentos ilustrativos de los distintos cambios de decisión son:

De aluminio a acero:

Para el uso que le voy a dar prefiero una bicicleta de aluminio, ya que pesa mucho menos. (PFI-NA-07, pretest). Las bicicletas de acero son más resistentes a golpes y caídas y, por tanto, más duraderas. Además, son más fáciles de reparar. (PFI-NA-07, postest).

De acero a aluminio:

Las bicicletas de acero son más baratas porque es un material más abundante y son más pesadas. Como no tengo mucha experiencia como ciclista, empezaría con una bicicleta más barata [...]. (PFI-NA-25, pretest). Estaba equivocado en mi razonamiento anterior. Las bicicletas de aluminio son más baratas y ligeras, por lo que presentan mayor ventaja. (PFI-NA-25, postest).

Argumentación

En el postest, la media de pruebas aumentó en ambos casos (de 1,35 a 2,83 pruebas/PFI-NA y de 2,21 a 2,68 pruebas/PFI-A), destacando el incremento en PFI-NA debido principalmente a que aportaron muy pocas pruebas en el pretest (tabla 4).

Los argumentos del postest destacan por su riqueza en una amplia tipología de pruebas independientemente de la decisión (Christenson et al., 2012; Ozturk y Yilmaz-Tuzun, 2017). Así, los argumentos de los PFI-NA siguen empleando mayoritariamente pruebas físicas (0,78 pruebas/PFI-NA), pero esta vez acompañadas de otras pruebas, económicas, químicas, o mecánicas, todas con una media por PFI de aproximadamente 0,50. En el caso de los PFI-A, las pruebas físicas siguen siendo mayoritarias (0,68 pruebas/PFI-A), junto a las mecánicas (0,64 pruebas/PFI-A), seguidas de las demás (más de 0,40 pruebas/PFI-A).

La calidad de las pruebas empeoró ligeramente en los PFI-NA (de 2,34 pruebas/PFI-NA a 2,00) y mejoró en los PFI-A (de 2,42 a 2,53). Dicho empeoramiento está relacionado con el empleo de la prueba económica incorrecta aludida, ya que 8 de los 13 PFI-NA que usaron dicha prueba (nivel 1) (tabla 4) incluían este error. Además, algunos PFI que expusieron pruebas económicas correctas en el pretest dieron por válida la información expuesta en el debate (véase ejemplo de PFI-NA-25). Otros PFI, pese a que dieron por válida esta prueba, mostraron sus dudas (*Elijo la bicicleta de aluminio. Tiene todas las prestaciones suficientes para un uso cotidiano siendo más barata. La debatiente a favor del acero ha dado argumentos que ponen en duda esta afirmación, por lo que no tengo claro si es realmente así. Tendría que informarme [...].*, PFI-A-36).

Por otro lado, la posibilidad de reflexionar sobre aspectos discutidos en el debate hizo que en el postest aumentaran ideas personales de los PFI-NA (de 0,25 puntos/PFI-NA a 0,52) mientras que para los PFI-A, disminuyó ligeramente (de 0,32 a 0,26).

Se detecta un progreso destacado en las justificaciones dadas (pretest: nivel 0-1, postest: nivel 1-2), encontrándose que tras el debate un elevado porcentaje de PFI fue capaz de dar justificaciones relacionando las pruebas con su conclusión (PFI-NA: 65,22 %; PFI-A: 84,21 %).

Asimismo, se avanza de forma muy importante en las conclusiones, puesto que el 100 % de los PFI ofreció una conclusión adecuada y precisa (nivel 2).

Los resultados muestran que las respuestas de los dos grupos de participantes siguen siendo bastante similares después del debate, considerándose en este caso decisiones bien argumentadas y basadas en conocimientos científicos en la mayoría de los casos (Sadler y Zeidler, 2005; Yacoubian, 2018). La prueba exacta de Fisher no mostró, como se ha indicado antes, diferencias significativas en decisiones ($\chi^2 = 0,71$; $p > 0,05$) ni ideas personales ($\chi^2 = 0,23$; $p > 0,05$), mientras que U de Mann-Whitney solo detectó diferencias estadísticamente significativas con un tamaño de efecto medio entre los dos grupos para la calidad de pruebas aportadas ($Z = -2,063$; $p = 0,039$, $r = 0,31$) a favor de PFI-A. Probable-

mente los argumentos del debate ayudaron a estos PFI-A a reforzar o completar algunas ideas que ya disponían de su formación afín de partida, por lo que mejoraron la calidad de las pruebas.

Impacto del debate

La prueba de McNemar no detectó diferencias estadísticamente significativas en las decisiones en ninguno de los casos analizados. Sin embargo, la actividad tuvo un impacto importante en la mejora de elaboración de argumentos en la totalidad de los PFI, pues la prueba de Wilcoxon detectó diferencias significativas a favor del postest en los tres elementos del argumento. Los avances se producen con un tamaño de efecto medio en el caso del número de pruebas ($Z = -3,785$; $p = 0,000$, $r = 0,41$), las pruebas económicas ($Z = -3,273$; $p = 0,001$, $r = 0,36$), otras pruebas ($Z = -2,841$; $p = 0,005$, $r = 0,31$) y las justificaciones ($Z = -3,593$; $p = 0,000$, $r = 0,39$), y con un tamaño de efecto pequeño para las pruebas mecánicas ($Z = -2,202$; $p = 0,028$, $r = 0,24$) y la conclusión ($Z = -2,236$; $p = 0,025$, $r = 0,24$).

El análisis sobre el posible impacto del debate según la afinidad del PFI reveló que las mejoras se producen principalmente en PFI-NA, donde se detectaron diferencias con un tamaño de efecto grande en el número de pruebas ($Z = -3,656$; $p = 0,000$, $r = 0,54$), con un tamaño de efecto medio para todos los tipos de pruebas excepto físicas (económicas y otras: $Z = -2,887$, $p = 0,004$, $r = 0,43$; químicas: $Z = -2,121$, $p = 0,034$, $r = 0,31$; mecánicas: $Z = -2,324$, $p = 0,020$, $r = 0,34$) y para las justificaciones ($Z = -2,385$; $p = 0,017$, $r = 0,35$) y con un efecto pequeño para la conclusión ($p = -2,000$; $p = 0,046$, $r = 0,29$).

Los PFI-A solo mejoraron en las justificaciones ($Z = -2,828$; $p = 0,005$, $r = 0,46$). Esto puede atribuirse a la afinidad en la formación previa de estos oyentes, lo que les otorgó conocimientos suficientes para presentar un mayor número y variedad de pruebas desde el principio en comparación con los PFI-NA, de acuerdo con Henderson et al. (2018) o Rodríguez-Mora et al. (2022), que indican que un bagaje de conocimientos científicos ayuda a construir buenos argumentos.

No obstante, es importante resaltar que esta actividad demostró ser efectiva en el desarrollo de argumentos (en lo que se refiere a la justificación) incluso para los PFI-A, lo que demuestra que, a pesar de que las pruebas (conocimientos) pueden ser adquiridas por distintos medios, la participación en esta actividad también contribuye a mejorar los argumentos expuestos.

En ninguno de los estudios se produjeron diferencias estadísticas en la calidad de las pruebas, ya que la media inicial por PFI era ya medio-alta y el error en las pruebas económicas expuestas en el postest no influyó lo suficiente.

CONCLUSIONES

De los resultados obtenidos se extraen las siguientes conclusiones para cada pregunta de investigación:

- *¿Qué decisiones adoptan los PFI sobre el problema tratado antes del debate y cómo las argumentan?* (P1)

Los dos grupos identificados en los participantes (PFI-A y PFI-NA) mostraron perfiles similares de respuestas en cuanto a la decisión adoptada y a los argumentos empleados. La decisión mayoritaria fue elegir la bicicleta de aluminio (83,33 %), respaldada principalmente por pruebas físicas. En todos los casos, la calidad de las pruebas fue medio-alta y no estuvieron focalizadas en ideas personales. Las conclusiones fueron adecuadas y precisas, siendo la ausencia de justificaciones el punto débil de los argumentos, por lo que las decisiones no se consideran totalmente argumentadas (Sadler y Zeidler, 2005; Yacoubian, 2018).

La única diferencia estadística detectada se encontró en que los PFI-A fueron capaces de aportar un mayor número de pruebas y otras pruebas (expertos) relacionadas con su formación.

- *¿Qué tipos de pruebas y justificaciones utilizan los PFI que intervienen en el debate?* (P2)
El debate estuvo equilibrado respecto a la calidad de los argumentos, las justificaciones, el uso de pruebas e ideas personales. Las pruebas mostraron diferentes dominios del uso del razonamiento informal (Christenson et al., 2012; Ozturk y Yilmaz-Tuzun, 2017; Wu, 2013), entre las que fueron predominantes las pruebas mecánicas y otras, de calidad de nivel 2-3 (medio-alto), acompañadas por ideas personales.
La preparación del debate ayudó a mejorar la dificultad detectada en la literatura asociada a la falta de conocimientos científicos para argumentar adecuadamente (Henderson et al., 2018), puesto que la debatiente sin formación previa sobre propiedades de los materiales fue capaz de mantener, en líneas generales, un debate equilibrado con la PFI-A, salvo al rebatir, cuando se detectó el uso de más ideas personales (Muñetón et al., 2017).
- *¿Cómo cambian, si lo hacen, las decisiones de los PFI y los argumentos que las apoyan después del debate?* (P3)
No se encontraron diferencias estadísticamente significativas entre la decisión inicial y final. El aluminio siguió siendo el material más elegido, aunque se detectaron algunos cambios de decisión en los dos sentidos, principalmente en el grupo de PFI-NA.
Las mejoras detectadas tras la actividad revelan la influencia del debate en los argumentos elaborados por los oyentes. Algunas evidencias son que todas las diferencias significativas sean a favor del posttest o el uso de una mayor variedad de pruebas (todos los tipos se emplearon en el debate) (Christenson et al., 2012) con buena calidad y con la aportación de justificaciones, donde se detectó una mejora importante, así como conclusiones en los niveles más altos de la rúbrica, salvo en la calidad de las pruebas. Probablemente esto se debe a que se partía de un valor de calidad alto en el pretest y en el debate se aportaron pruebas económicas confusas. El debate no influyó en la incorporación de ideas personales, donde no hubo diferencias estadísticas.
Las mejoras son más relevantes en PFI-NA que PFI-A, ya que estos últimos, al disponer de un bagaje sobre el tema, solo mejoraron en las justificaciones, lo que revela la importancia de disponer de conocimientos científicos para argumentar adecuadamente (Henderson et al., 2018). El impacto del debate en estos avances fue importante, ya que en él se mostró un amplio conjunto de pruebas bien justificadas. Finalmente, las decisiones adoptadas tras el debate se pueden considerar bien argumentadas y basadas en conocimientos científicos en la mayoría de los casos (Sadler y Zeidler, 2005; Jiménez-Aleixandre, 2010).

LIMITACIONES E IMPLICACIONES EDUCATIVAS

Consideramos que la propuesta formativa se podría completar incluyendo una actividad final grupal que tenga por objetivo mostrar a los PFI el análisis de los argumentos utilizados para que sean conscientes de sus carencias y mejoras y puedan contribuir de esta forma a mejorar su argumentación. En este análisis, resulta de interés centrarnos también en la contraargumentación. Si bien es cierto que en la investigación presentada los argumentos analizados incluyen también contraargumentos, aunque no de manera separada, es fundamental mostrar al alumnado ejemplos de refutaciones como factor importante para conseguir argumentos de calidad (Jiménez-Aleixandre, 2010).

Otra limitación reside en que durante la actividad no aparecieron pruebas relacionadas con la sostenibilidad o el impacto ambiental de la producción de los diferentes materiales, así como su gestión posterior, aspecto clave para mejorar la conciencia ambiental a partir de este problema de la vida diaria (Hadjichambis et al., 2020). Por tanto, habría que abordar de manera explícita estos aspectos (Moreno-Fontiveros et al., 2022).

Finalmente, sería conveniente mostrar a los PFI la metodología y los resultados de este estudio y que sean conscientes de los beneficios que tienen para los oyentes y participantes para mejorar sus habilidades de toma de decisiones y argumentación a partir de problemas de la vida diaria (Moreno-Fon-tiveros et al., 2022), basándose en conocimientos científicos y tecnológicos (Sadler y Zeidler, 2005) y minimizando el uso de creencias, valores u opiniones personales (von Winterfeldt, 2013). Esta toma de conciencia puede ser una de las razones que les puede impulsar a utilizar este tipo de actividades en su futura práctica profesional.

AGRADECIMIENTOS

Este trabajo forma parte de los siguientes proyectos:

- Proyecto I+D+i «Ciudadanos con pensamiento crítico: Un desafío para el profesorado en la enseñanza de las ciencias» (PID2019-105765GA-I00) financiado por MCIN/AEI/10.13039/501100011033. El estudio se realizó de acuerdo con el protocolo aprobado por el Comité Ético de Experimentación de la Universidad de Málaga (CEUMA), referencia 31-2022-H.
- Beca de Iniciación a la Investigación, modalidad B (Máster), del I Plan Propio de Investigación, Transferencia y Divulgación Científica de la Universidad de Málaga, concedida a la primera autora de este trabajo en el curso 2021-2022.
- Proyecto de Innovación Educativa PIE22-184 del Grupo permanente de Innovación en Educación Crítica (EDUCRIT), financiado por la Universidad de Málaga, convocatoria INNOVA22 para bienio 2022-23.

REFERENCIAS

- Acevedo, J. A. (2006). Relevancia de los factores no-epistémicos en la percepción pública de los asuntos tecnocientíficos. *Revista Eureka sobre Enseñanza y Divulgación de las Ciencias*, 3(3), 370-391. http://dx.doi.org/10.25267/Rev_Eureka_ensen_divulg_cienc.2006.v3.i3.03
- Bailin, S. (2002). Critical thinking and science education. *Science & Education*, 11(4), 361-375. <https://doi.org/10.1023/A:1016042608621>
- Bargiela, I. M., Blanco, P. y Puig, B. (2022). Las preguntas para la indagación y activación de pensamiento crítico en educación infantil. *Enseñanza de las Ciencias*, 40(3), 11-28. <https://doi.org/10.5565/rev/ensciencias.5470>
- Blanco, H. A. y Forero, D. L. (2017). *Comportamiento del consumidor frente a la decisión de compra en artículos deportivos relacionados con el uso de bicicleta en la ciudad de Bogotá* [Tesis de doctorado]. Universidad Externado de Colombia.
- Blanco, A., España, E. y Franco-Mariscal, A. J. (2017). Estrategias didácticas para el desarrollo del pensamiento crítico en el aula de ciencias. *Ápice, Revista de Educación Científica*, 1(1), 107-115. <https://doi.org/10.17979/arec.2017.1.1.2004>
- Bravo, B. y Jiménez-Aleixandre, M. P. (2018). Developing an Initial Learning Progression for the Use of Evidence in Decision-Making Contexts. *International Journal of Science and Mathematics Education*, 16(4), 619-638. <https://doi.org/10.1007/s10763-017-9803-9>

- Cano-Iglesias, M. J., Franco-Mariscal, A. J. y Blanco-López, A. (2021). Secuencia de actividades de argumentación para estudiantes de ingenierías industriales. En D. Cebrián-Robles, A. J. Franco-Mariscal, T. Lupión-Cobos, M. C. Acebal-Expósito y A. Blanco-López (Coords.), *Enseñanza de las ciencias y problemas relevantes de la ciudadanía. Transferencia al aula* (pp. 153-172). Graó.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Erlbaum.
- Christenson, N., Chang-Rundgren, S. N. y Höglund, H. O. (2012). Using the SEE-SEP Model to Analyze Upper Secondary Students' Use of Supporting Reasons in Arguing Socioscientific Issues. *Journal of Science Education and Technology*, 21, 342-352.
<https://doi.org/10.1007/s10956-011-9328-x>
- Cruz-Lorite, I. M., Cebrián, D., Acebal, M. C. y Evagorou, M. (2023). Analysis of the Informal Reasoning Modes of Preservice Primary Teachers When Arguing about a Socio-Scientific Issue on Nuclear Power during a Role Play. *Sustainability*, 15, 4291.
<https://doi.org/10.3390/su15054291>
- Duschl, R. A., Schweingruber, H. A. y Shouse, A.W. (Eds.) (2007). *Taking science to school: Learning and teaching science in grades K-8*. National Research Council.
- Erduran, S., Simon, S. y Osborne, J. (2004). TAPping into argumentation: Developments in the application of Toulmin's Argument Pattern for studying science discourse. *Science Education*, 88(6), 915-933.
<https://doi.org/10.1002/sce.20012>
- Fang, S. C., Hsu, Y. S. y Lin, S. S. (2019). Conceptualizing socioscientific decision making from a review of research in science education. *International Journal Science and Mathematics Education*, 17, 427-448.
<https://doi.org/10.1007/s10763-018-9890-2>
- Girón, J. R. y Lupión, T. (2022). Influencia de la publicidad en los argumentos de adolescentes sobre consumo alimentario. *Enseñanza de las Ciencias*, 40(2), 167-192.
<https://doi.org/10.5565/rev/ensciencias>
- Grosick, T. L., Talbert, C., Myers, M. J. y Angelo, R. (2013). Assessing the landscape: Body image values and attitudes among middle school boys and girls. *American Journal of Health Education*, 44(1), 41-52.
<https://doi.org/10.1080/19325037.2012.749682>
- Jenkins, L. (2011). Using citizen science beyond teaching science content: A strategy for making science relevant to students' lives. *Cultural Studies of Science Education*, 6(2), 501-508.
- Jiménez, L. y Otero, J. (2019). La educación científica frente al pensamiento anticrítico en la vida diaria. *Enseñanza de las Ciencias*, 37(1), 117-135.
<https://doi.org/10.5565/rev/ensciencias.2608>
- Jiménez-Aleixandre, M. P. (2010). *10 ideas clave. Competencias en argumentación y uso de pruebas*. Graó.
- Hadjichambis, A. C., Reis, P., Paraskeva-Hadjichambi, D., Činčera, J., Boeve-de Pauw, J., Gericke, N., y Knippels, M. C. (Eds.) (2020). *Conceptualizing Environmental Citizenship for 21st Century Education*. Springer.
<https://doi.org/10.1007/978-3-030-0249-1>
- Henderson, J. B., McNeill, K. L., González-Howard, M., Close, K. y Evans, M. (2018). Key challenges and future directions for educational research on scientific argumentation. *Journal of Research in Science Teaching*, 55(1), 5-18.
<https://doi.org/10.1002/tea.21412>
- López, G. (2012). Pensamiento crítico en el aula. *Docencia e Investigación*, 37(22), 41-60.

- Lupión, T., López, R. y Blanco, A. (2017). What do science teachers think about developing scientific competences through context-based teaching? A case study. *International Journal of Science Education*, 39(7), 937-963.
<https://doi.org/10.1080/09500693.2017.1310412>
- Martín, M. C. y Salamanca, A. B. (2007). El muestreo en la investigación cualitativa. *Nure Investigación*, 27, 1-4.
- Moreno-Fontiveros, G., Blanco, A. y España, E. (2015). Importancia del ahorro energético en la decisión de comprar un coche. Un estudio en 3º de ESO. *Alambique, Didáctica de las Ciencias Experimentales*, 80, 29-37.
- Moreno-Fontiveros, G., Cebrián, D., Blanco, A. y España, E. (2022). Decisiones de estudiantes de 14/15 años en una propuesta didáctica sobre la compra de un coche. *Enseñanza de las Ciencias*, 40(1), 199-219.
<https://doi.org/10.5565/rev/ensciencias.3292>
- Muñetón, G., Ruiz, A. F., y Loaiza, O. L. (2017). Toma de decisiones: Explicaciones desde la ciencia aplicada del comportamiento. *Espacios*, 38(13), 1-12.
- Osborne, J. (2001). Promoting Argument in the Science Classroom: A Rhetorical Perspective. *Canadian Journal Science, Mathematics and Technology Education*, 3(1), 271-290.
<https://doi.org/10.1080/14926150109556470>
- Osborne, J. (2014). Teaching critical thinking. New directions in science education? *School Science Review*, 352, 53-62.
- Osborne, J., Henderson, J. B., MacPherson, A., Szu, E., Wild, A. y Yao, S. (2016). The development and validation of a learning progression for argumentation in science. *Journal Research in Science Teaching*, 53(6), 821-846.
<https://doi.org/10.1002/tea.21316>
- Ozturk, N. y Yilmaz-Tuzun, O. (2017). Preservice science teachers' epistemological beliefs and informal reasoning regarding socio-scientific issues. *Research in Science Education*, 47, 1275-1304.
<https://doi.org/10.1007/s11165-016-9548-4>
- Prieto, T., España, E. y Martín, C. (2012). Algunas cuestiones relevantes en la enseñanza de las ciencias desde una perspectiva Ciencia-Tecnología-Sociedad. *Revista Eureka sobre Enseñanza y Divulgación de las Ciencias*, 9(1), 71-77.
http://dx.doi.org/10.25267/Rev_Eureka_ensen_divulg_cienc.2012.v9.i1.05
- Rivera, C., Arellano, C. y Molero, A. (2009). *Conducta del consumidor: Estrategias y políticas aplicadas al marketing* (2.ª ed.). ESIC Editorial.
- Rodríguez, R. (2012). El debate como estrategia de innovación docente. Experiencias en filosofía del derecho y teoría de la cultura. *UPO Innova: Revista de Innovación Docente*, 1, 493-503.
- Rodríguez-Mora, F., Cebrián, D. y Blanco, A. (2022). An Assessment Using Rubrics and the Rasch Model of 14/15-Year-Old Students' Difficulties in Arguing About Bottled Water Consumption. *Research in Science Education*, 52, 1075-1091.
<https://doi.org/10.1007/s11165-020-09985-z>
- Sadler, T. D. y Zeidler, D. L. (2005). Patterns of informal reasoning in the context of socioscientific decision making. *Journal of Research in Science Teaching*, 42(1), 112-138.
<https://doi.org/10.1002/tea.20042>
- Simonneaux, L. (2001). Role-play or debate to promote students' argumentation and justification on an issue in animal transgenesis. *International Journal Science Education*, 23(9), 903-927.
<https://doi.org/10.1080/09500690010016076>
- Siribunnam, S., Nuangchalerm, P. y Jansawang, N. (2014). Socio-Scientific Decision Making in the Science Classroom. *Online Submission*, 5(4), 1777-1782.

- Solbes, J. y Torres, N. (2012). Análisis de las competencias de pensamiento crítico desde el abordaje de las cuestiones sociocientíficas: un estudio en el ámbito universitario. *Didáctica de las Ciencias Experimentales y Sociales*, 26, 247-269.
<https://doi.org/10.7203/dces.26.1928>
- Torres, N. y Solbes, J. (2016). Contribuciones de una intervención didáctica usando cuestiones sociocientíficas para desarrollar el pensamiento crítico. *Enseñanza de las Ciencias*, 34(2), 43-65.
<http://dx.doi.org/10.5565/rev/ensciencias.1638>
- Toulmin, S. E. (2003). *The uses of argument* (3.^a ed.). Cambridge University Press.
- Villani, S. (2001). Impact of Media on Children and Adolescents: A 10-Year Review of the Research. *Journal of the American Academy of Child and Adolescent Psychiatry*, 40(4), 392-401.
<https://doi.org/10.1097/00004583-200104000-00007>
- von Winterfeldt, D. (2013). Bridging the gap between science and decision making. *Proceedings of the National Academy of Sciences of the United States of America*, 110 (Supp. 3), 14055-14061.
<https://doi.org/10.1073/pnas.1213532110>
- Yacoubian, H. A. (2018). Scientific literacy for democratic decision-making. *International Journal of Science Education*, 40(3), 308-327.
<https://doi.org/10.1080/09500693.2017.1420266>
- Wang, H. H., Chen, H. T., Lin, H. S., Huang, Y. N. y Hong, Z. R. (2017). Longitudinal study of a cooperation-driven, socio-scientific issue intervention on promoting students' critical thinking and self-regulation in learning science. *International Journal Science Education*, 39(15), 2002-2026.
<https://doi.org/10.1080/09500693.2017.1357087>
- Wu, Y. T. (2013). University students' knowledge structures and informal reasoning on the use of genetically modified foods: multi-dimensional analyses. *Research in Science Education*, 43, 1873-1890.
<https://doi.org/10.1007/s11165-012-9343-9>

Impact of a Socio-Scientific Debate on the Argumentative and Decision-Making Skills of Preservice Secondary Teachers

Pilar Bernal-Herrera, María José Cano-Iglesias, Antonio Joaquín Franco-Mariscal y Ángel Blanco-López
Didáctica de las Ciencias Experimentales. Universidad de Málaga. Málaga. España
pilar_9@uma.es; mjcano@uma.es, anjoa@uma.es, ablancol@uma.es

This research explores the impact that a debate about the choice of an everyday product has on argumentation and decision-making skills, which are important dimensions of critical thinking. The study involves 45 students (59.5 % females and 40.5 % males) specializing in sciences in the master's degree in Secondary Education at the University of Malaga (Spain). This case involves a choice between materials (steel or aluminum) when purchasing a bicycle. The activity consists of a brief debate, featuring a presenter and two debaters taking opposing stances, while the rest of the students act as listeners. The debate follows a structure, beginning with an impartial presentation of the issue by the presenter (three minutes), followed by one-minute opening statements from each debater (prepared in advance), a five-minute discussion, and concluding with one-minute closing remarks.

The aim is to comprehend the decisions taken by the master's students before the debate and to analyze their argumentation. Additionally, we examine whether and how these decisions change after the debate and what types of evidence and justifications are employed by the participants in this discussion.

The data collection instruments included a pre- and post-test concerning the decisions and arguments of the listeners, as well as the preparation reports and audio recordings of the debate. An analysis of the decisions taken before and after the debate and their potential changes was conducted, along with an argument analysis, which included a qualitative study to define response categories and a statistical treatment using non-parametric tests. All arguments, whether written or oral, were assessed using a rubric based on the simplified Toulmin model, identifying evidence, justifications, and conclusions. For the evidence, the number, type, and quality were analyzed, and the inclusion of personal ideas was also considered.

Students were divided into two groups based on their prior education: one group had studied degrees that included training in material properties (related degrees such as Chemistry, Chemical Engineering, etc.), while the other group had degrees that did not encompass this training (unrelated grades such as Biology, Veterinary science, etc.).

Students with both related and unrelated degrees exhibited similar initial response patterns in terms of their choices and the arguments they employed. Most of them chose the aluminum bicycle (83.33 %), supported by physical evidence. In all instances, the quality of the evidence was moderately to highly reliable, with minimal emphasis on personal opinions. Conclusions were appropriate and precise, although the weakness in the arguments stem from the absence of justifications.

The debate was balanced in terms of argument quality, justifications, use of evidence, and personal ideas. The evidence showed different domains of informal reasoning, with mechanical evidence being predominant alongside personal ideas. Debate preparation helped to mitigate the challenges associated with a lack of scientific knowledge in presenting arguments effectively.

No statistically significant differences were detected between the initial and final decisions. Aluminum remained the most chosen material, although some decision changes were observed, mainly among students from unrelated degree programs. The improvements identified after the activity highlight the impact of the debate on the arguments presented by the listeners. Some evidence includes a broader use of high-quality evidence in the post-test, as well as the inclusion of justifications and conclusions at the highest levels of the rubric.

The progress is more relevant among students from unrelated degrees than among those from related programs, as the latter, having a background in the subject, only improved in justifications. The impact of the debate was significant, demonstrating a wide range of well-justified evidence. Finally, the decisions after the debate can be considered well-argued and grounded in scientific knowledge in most cases.