

Conflictos semióticos relacionados con el intervalo de confianza en estudiantes de Bachillerato e Ingeniería

Semiotic conflicts related to confidence interval in high school and engineering students

Antonio Francisco Roldán-López Hierro

Carmen Batanero

Rocío Álvarez-Arroyo

Resumo: La importancia que hoy día recibe el intervalo de confianza se refleja en la inclusión del tema en el Bachillerato de Ciencias Sociales y en varios estudios universitarios, incluidos los de ingeniería. La finalidad del trabajo fue comparar la comprensión de las propiedades esenciales del intervalo en estos dos tipos de estudiantes. Para ello se analizan las respuestas de 58 estudiantes de Bachillerato y 37 de ingeniería a un cuestionario que consta de seis ítems de opción múltiple. Los resultados muestran interpretaciones bayesianas del intervalo, falta de comprensión de la forma en que se relaciona su amplitud con el tamaño de la muestra o la varianza poblacional, y desconocimiento de la finalidad de la estimación por intervalos, siendo los resultados algo mejores en el segundo grupo de estudiantes.

Palabras clave: Intervalo de confianza. Comprensión. Estudiantes de Bachillerato e Ingeniería.

Abstract: The relevance that the confidence interval receives today is reflected in the inclusion of the subject in the Bachelor of Social Sciences and in university studies, including engineering. The purpose of the work was to compare the understanding of the essential properties of the interval in these two types of students. To achieve this goal, the responses of 58 high school and 37 engineering students to a questionnaire consisting of six multiple-choice items are analysed. The results show Bayesian interpretations of the interval, lack of understanding of the way in which its amplitude is related to the sample size or population variance, and ignorance of the purpose of the estimation by intervals, the results being somewhat better in the engineering students.

Keywords: Confidence interval. Understanding. Engineering and High School students.

1 Introducción

El intervalo de confianza es uno de los procedimientos inferenciales más utilizados en la actualidad para complementar el contraste de hipótesis, debido a las recomendaciones de diferentes investigadores dirigidas a mejorar el uso de la inferencia por parte de los investigadores (THOMPSON, 2007; CUMMING, 2012; YAREMKO *et al.*, 2013). Dicho intervalo puede ser

Antonio Francisco Roldán-López Hierro
Doctor en Ciencias Matemáticas. Profesor de la Facultad de Ciencias de la Universidad de Granada. Granada, España.

 orcid.org/0000-0002-6956-4328

 aroldan@ugr.es

Carmen Batanero
Doctora en Matemáticas. Profesora Catedrática de Didáctica de la Matemática en la Facultad de Ciencias de la Educación de la Universidad de Granada. Granada, España.

 orcid.org/0000-0002-4189-7139

 batanero@ugr.es

Rocío Álvarez-Arroyo
Doctora en Ingeniería Civil. Profesora la Facultad de Ciencias de la Educación de la Universidad de Granada. Granada, España.

 orcid.org/0000-0002-3201-8542

 rocioaarroyo@ugr.es

Recibido el 25/01/2020
Aceptado el 28/02/2020
Publicado el 07/04/2020

construido en la actualidad con diferentes metodologías que no sólo se diferencian en los procedimientos matemáticos, sino en sus fundamentos filosóficos, aunque todas tienen como finalidad tener en cuenta la variabilidad muestral al estimar un parámetro poblacional θ (MOREY *et al.*, 2016).

Dicho parámetro, por ejemplo, la media o proporción de un cierto carácter en una población, es desconocido y el interés en estimarlo es porque determina la distribución de la variable en la población. Cuando no es posible tomar toda la población, se selecciona una muestra aleatoria de la misma a partir de la cual se calcula un estimador del parámetro $\hat{\theta}$ que puede variar de una muestra a otra. Así, para el caso de la media poblacional, se usa como estimador la media muestral y siguiendo a Neyman (1937) se utiliza la desviación típica $S_{\hat{\theta}}$ de dicho estimador para expresar la precisión de la estimación mediante un intervalo de la forma:

$$EI = \hat{\theta} - k_1 S_{\hat{\theta}} \text{ y } ES = \hat{\theta} + k_2 S_{\hat{\theta}}$$

Los extremos del intervalo (EI y ES) se calculan en función del coeficiente de confianza $1 - \alpha$, siendo $0 < \alpha < 1$, escogido por el investigador, usualmente 0,95 o 0,99 (MAYO, 1981). Una vez calculado el intervalo, se admite que el valor verdadero del parámetro se encuentra en el intervalo determinado. El valor $1 - \alpha$ o confianza en la estimación, según Mayo (1981) se refiere al porcentaje de intervalos construidos a partir de diferentes muestras del mismo tamaño de la población que cubrirán el parámetro y no deben interpretarse como probabilidad de que el verdadero valor del parámetro θ esté situado dentro de un intervalo particular, lo que sería una interpretación bayesiana del mismo.

La importancia de que los estudiantes comprendan los fundamentos de la construcción y la filosofía del intervalo de confianza ha sido reconocida tanto en el Bachillerato como en los cursos universitarios. Así, en el currículo español para el Bachillerato de Ciencias Sociales (MECD, 2015) se sugiere como tema de estudio la estimación por intervalos de confianza; relación entre confianza, error y tamaño muestral; intervalo de confianza para la media poblacional de una distribución normal con desviación típica conocida; intervalo de confianza para la media poblacional de una distribución de modelo desconocido y para la proporción en el caso de muestras grandes. Aunque el Bachillerato de Ciencias no incluye temas de inferencia, el intervalo de confianza, con contenidos similares a los descritos, se incluye en los cursos universitarios de estadística y en particular en las carreras de ingeniería. La importancia en el Bachillerato se acentúa finalmente por el hecho de que las pruebas de acceso a la universidad para el Bachillerato

de Ciencias Sociales incluyen habitualmente un problema de intervalos de confianza (LÓPEZ-MARTÍN *et al.*, 2016)

Es importante, entonces, asegurar la comprensión del tema por parte de los estudiantes, pues diferentes autores indican la dificultad generalizada de la inferencia, incluso para investigadores y profesores (CASTRO SOTOS *et al.*, 2007; HARRADINE, BATANERO y ROSSMAN, 2011). Estas investigaciones se han concentrado principalmente en el contraste de hipótesis y en el uso de la inferencia en las publicaciones científicas, con escasos estudios sobre la comprensión de los estudiantes españoles.

Con la finalidad de aportar información al respecto, el objetivo de este estudio exploratorio es proporcionar información sobre la comprensión de los estudiantes de las principales propiedades del intervalo de confianza y más concretamente comparar dicha comprensión en estudiantes del Bachillerato de Ciencias Sociales y estudiantes de ingeniería después de haber estudiado el tema. A continuación, se exponen los fundamentos, metodología y resultados de la investigación.

2 Fundamentos

Para fundamentar el trabajo utilizamos las ideas de comprensión y conflicto semiótico tomados del enfoque ontosemiótico (GODINO, 1996, 2002; GODINO, BATANERO y FONT, 2007). Además, nos apoyamos en los trabajos previos de investigación sobre la comprensión del intervalo de confianza recogidos en los antecedentes de este trabajo.

2.1 Enfoque ontosemiótico

El enfoque ontosemiótico es un marco teórico que trata de explicar el aprendizaje y la comprensión de la matemática que, según los autores, surge de las prácticas realizadas al resolver problemas matemáticos. En dicho marco teórico se diferencia entre significado institucional y personal de los objetos matemáticos, el primero de los cuales engloba las prácticas que dentro de una institución se realizan y el segundo las prácticas prototípicas de cada persona respecto a dicho objeto. Serían, por ejemplo, instituciones diferentes en este marco el grupo de estadísticos que siguen la metodología frecuencial o la bayesiana porque para resolver problemas asociados al mismo objeto (en este caso, la estimación por intervalos), se realizan prácticas diferentes en cada

institución. Igualmente, el significado personal que un estudiante asigna al intervalo de confianza puede ser diferente del significado institucional en el nivel de estudios que cursa, por ejemplo, si asigna al intervalo de confianza propiedades incorrectas desde el punto de vista matemático.

Godino (1996) concibe la comprensión de un objeto matemático por parte de un estudiante como la coincidencia entre los significados institucional y personal sobre dicho objeto. Dicha comprensión se adquiere gradualmente a lo largo de los estudios y puede ser parcial, pues puede haber un grado de acuerdo en una parte del significado institucional del objeto y un desacuerdo en otra parte. En este sentido, en nuestro estudio se trata de evaluar el acuerdo o desacuerdo de los estudiantes de la muestra con la definición y de las propiedades básicas del intervalo de confianza.

En el enfoque ontosemiótico, además de los problemas que dan sentido a los objetos matemáticos, se distinguen diferentes tipos de objetos que intervienen en las prácticas matemáticas: el lenguaje matemático (verbal, simbólico, gráfico), los conceptos y propiedades de los mismos, procedimientos de cálculo y razonamiento matemático. Todos ellos intervienen en el trabajo matemático y se relacionan entre sí formando funciones semióticas (o correspondencias) que constan de una expresión o antecedente (el signo), el contenido de la expresión (significado de tal signo, lo representado) y el criterio (o regla de correspondencia) que permite expresar la relación entre las dos componentes anteriores. Por ejemplo, utilizamos " $z_{\alpha/2}$ " como signo para representar el valor crítico (contenido) asociado al nivel de confianza fijado, quedando establecido un vínculo (criterio) entre ellos que el alumnado debe conocer.

Debido a la complejidad del lenguaje y la práctica matemática, en ocasiones, según los autores, se produce un conflicto semiótico cuando existe una discrepancia entre el significado que en la enseñanza se ha establecido de la función semiótica (significado institucional) y el que manifiesta la persona encargada de su interpretación (significado personal). A continuación, describimos algunos de estos conflictos que han sido descritos como errores en investigaciones previas.

2.2 Antecedentes

Una parte de investigaciones sobre la comprensión de los intervalos de confianza se han realizado tomando como sujetos a investigadores en psicología o educación, bien a partir de entrevistas o encuestas a los mismos, o mediante el análisis de su interpretación de intervalos de

confianza en artículos publicados en revistas de investigación de impacto.

Uno de los estudios más citados fue el de Cumming, William y Fidler (2004), quienes entrevistaron a 134 autores de artículos publicados en revistas de impacto en los que habían utilizado intervalos de confianza para ver si entendían el concepto de replicación. Tras proporcionarles un intervalo de confianza para la media de una población, construido a un nivel de confianza del 95%, se les preguntó sobre los valores que esperaban de la media muestral si se repitiese el mismo experimento en un gran número de ocasiones. Más de tres cuartas partes de los investigadores participantes en el estudio respondieron que la media volvería a caer en el intervalo proporcionado en el 95% de las veces, lo que es una interpretación propia del método de remuestreo y no del método frecuencial de construcción de intervalos.

Otro estudio realizado esta vez con investigadores y estudiantes es el de Behar (2001) que proporcionó un cuestionario a 47 investigadores y 297 estudiantes de ingeniería en Colombia. Uno de los conflictos identificado por el autor en el 29% de los investigadores y el 50% de estudiantes es interpretar el nivel de confianza como porcentaje de valores de la variable que cae en el intervalo. Además, el 40% de los investigadores y 50% de los estudiantes interpretaron el nivel de confianza como probabilidad de que el intervalo contenga al verdadero valor del parámetro poblacional, lo que sería una interpretación bayesiana y no frecuencial del intervalo. La interpretación frecuencial correcta es que algunos intervalos no contienen al verdadero valor del parámetro, pero al repetir muchas veces el experimento, el porcentaje de intervalos que contienen al parámetro poblacional se acerca al nivel de confianza. Otros errores fueron creer que al aumentar el nivel de confianza dejando invariantes los demás factores se obtienen intervalos más estrechos e incluso suponer que existe proporcionalidad directa entre la anchura del intervalo y el tamaño de la muestra.

Otros conflictos que aparecen al interpretar el intervalo de confianza por parte de los estudiantes son identificados por Fidler y Cumming (2005) en una muestra de 180 estudiantes de psicología: un 38% interpretaban el intervalo como un conjunto de valores razonables para la media muestral, y un 19% confundía el rango de la variable aleatoria con el intervalo de confianza. Por otro lado, un 20% pensaba que la anchura del intervalo aumentaría con el tamaño muestral, un 29% consideró que no variaría y un 36% no fue capaz de describir el tipo de relación entre la anchura y el tamaño de la muestra.

Olivo y Batanero (2007) estudiaron la comprensión del intervalo de confianza en 48

estudiantes de ingeniería teniendo en cuenta los contenidos de un análisis previo del tema en libros de texto universitarios utilizados por los estudiantes. Aunque la mayor parte de los estudiantes respondieron correctamente muchas preguntas, los autores también describen los siguientes conflictos semióticos de algunos de ellos:

- El 27% de los alumnos dio una interpretación bayesiana al intervalo, identificando el nivel de confianza como la probabilidad de que el intervalo contenga al parámetro poblacional, mientras otro 21% interpreta que el intervalo contiene un porcentaje de valores de la variable, en vez de comprender que es una estimación del parámetro.
- La relación entre el tamaño de la muestra y la precisión del intervalo sólo es comprendida por el 25% de los alumnos, mientras que el 18% sugirió que el nivel de confianza no afecta a la anchura del intervalo.

Olivo, Batanero y Díaz (2008) analizan la comprensión del intervalo de confianza en 252 estudiantes de ingeniería, de los cuales el 25% mostró confusión entre las ideas de estadístico y parámetro, pensando que el intervalo se construye para predecir los valores de la media muestral, lo que en realidad sería la interpretación propia de los intervalos construidos con el método de remuestreo. Otro 33% no comprendía la relación entre ancho del intervalo y tamaño de muestra o coeficiente de confianza y un 32% de estudiantes que habían calculado correctamente el intervalo le daban una interpretación bayesiana.

Estos conflictos de interpretación también aparecen en ocasiones en los estudiantes que se preparan para ser profesor de matemáticas, como se muestra en la investigación de López-Martín *et al.* (2019a), quienes evaluaron el conocimiento matemático del intervalo de confianza de 70 futuros profesores de matemáticas de secundaria a los que se pidió construir e interpretar un intervalo de confianza. Mientras sólo un 28% de los participantes dio una interpretación correcta, el 11% afirmó que el intervalo debía contener, con seguridad, el verdadero valor del parámetro (visión determinista), error previamente indicado por Olivo (2008). Un 17.8% dio la interpretación bayesiana, conflicto ya detectado por Behar (2001) y Olivo (2008); un 41.1% no aportó ninguna interpretación (ni correcta ni incorrecta) sobre el significado del intervalo de confianza; y un estudiante pensó que el intervalo sirve para estimar el estadístico muestral y no el poblacional.

3 Metodología

3.1 Muestra participante

La muestra estuvo formada por dos grupos diferentes de estudiantes: El primero estuvo formado por 58 estudiantes del último curso de Bachillerato de la especialidad de Ciencias Sociales, que provenían de dos institutos diferentes: uno situado en la ciudad de Granada, del que participaron 24 alumnos matriculados en dos grupos diferentes, mientras que el otro centro está situado en la periferia de la capital, donde participaron 34 estudiantes también de dos grupos. El segundo grupo consistió en 35 estudiantes universitarios de ingeniería de la Universidad de Granada, 23 de los cuales estaban cursando la asignatura de “Estadística y Optimización” de primer curso del Grado en Ingeniería de Telecomunicaciones y otros 12 la asignatura de “Estadística” del primer curso del Grado en Ingeniería Informática.

Como acabamos de poner de manifiesto, se trata de una muestra intencional por lo que nuestro estudio es de tipo exploratorio, de manera que no pretendemos extender las conclusiones obtenidas a una muestra o a una población más amplia.

Otra diferencia fundamental respecto de los dos grupos es que los estudiantes preuniversitarios habían cursado inferencia como parte del Bachillerato Ciencias Sociales, mientras que el alumnado universitario de ingeniería provenía del Bachillerato de Ciencias, donde no se estudia la inferencia estadística. En consecuencia, aun cuando sus niveles educativos son claramente diferentes, todos los estudiantes participantes en el estudio se estaban enfrentando por primera vez al estudio del tema de la inferencia estadística.

Todos los cuestionarios fueron completados a lo largo del mes de mayo de 2019, cuando ambos grupos de estudiantes habían finalizado el estudio de la inferencia y se preparaban para las pruebas de evaluación finales de sus correspondientes asignaturas. Se eligió este periodo temporal por ser un momento en que los estudiantes ya habían preparado el tema. El cuestionario se completó dentro de la hora de clase como una actividad que hiciera reflexionar a los participantes sobre su grado de comprensión sobre los intervalos de confianza.

3.2 Descripción del cuestionario

El cuestionario utilizado se construyó para esta investigación partiendo del análisis de otros cuestionarios empleados en investigaciones previas y consta de seis ítems de opción

múltiple. Los ítems 1 a 3 están adaptados de otros utilizados por Olivo (2008), el 4 de Behar (2001) y los ítems 5 y 6 son de elaboración propia. En la Tabla 1 mostramos un resumen del contenido evaluado por el cuestionario que, como se puede observar, evalúa un amplio abanico de propiedades del intervalo de confianza junto con algunos de los conflictos previamente detectados en la investigación.

Tabla 1. Contenido evaluado por el cuestionario

Contenido	I1	I2	I3	I4	I5	I6
Definición del intervalo de confianza	x					
Porcentaje de intervalos que cubre el valor del parámetro	x					
Extremos aleatorios en el intervalo	x					
Confusión entre media muestral y poblacional	x					
Considerar fijos los extremos del intervalo	x					
Confusión entre confianza y probabilidad (interpretación bayesiana)	x					
Relación del ancho del intervalo y tamaño de la muestra		x	x			
Confusión entre confianza y amplitud		x	x			
Relación entre precisión y confianza		x	x			
Fórmula de la desviación típica de la distribución muestral		x				
Relación entre amplitud del intervalo y varianza de la población				x		
Relación del intervalo con la media muestral					x	
Algunos intervalos no cubren el parámetro					x	
El ancho del intervalo no depende de la media muestral					x	
El intervalo sirve para estimar valores desconocidos del parámetro						x
Pensar que el intervalo siempre contiene al parámetro						x

Fuente: Elaboración de los Autores

4 Resultados

A continuación, se analiza el contenido de cada uno de los seis ítems que componen la primera parte del cuestionario. Se resaltan en **negrita** las respuestas correctas. Para analizar las respuestas de los estudiantes a las preguntas incluidas en el cuestionario, en las siguientes secciones presentamos en diferentes tablas las frecuencias con las que los participantes ha escogido cada opción y los porcentajes que éstas representan, diferenciando entre el alumnado

preuniversitario y el universitario. Obsérvese que la suma de respuestas puede no coincidir con el total de alumnos encuestados ni la suma de porcentajes corresponder al 100% ya que, cada estudiante, podía elegir una o varias respuestas.

4.1 Resultados en el ítem 1

Ítem 1. Se calcula un intervalo de confianza del 90% para la media μ de una población. ¿Cuál o cuáles de las siguientes afirmaciones son verdaderas?

- (a) Si se toman muchas muestras, la media muestral \bar{x} caerá dentro del intervalo de confianza aproximadamente el 90% de las veces.
- (b) El intervalo de confianza es un intervalo de valores calculado a partir de los datos de la muestra. En el 90% de las muestras de una población, el intervalo calculado contiene a la media.**
- (c) El intervalo de confianza es un intervalo con extremos constantes, dentro del cual cae la media poblacional el 90% de las veces.
- (d) La probabilidad de que μ caiga dentro de un intervalo de confianza calculado a partir de una muestra es de 0,90.

Mediante este ítem se pretende evaluar la comprensión de la definición del intervalo de confianza y detectar en los estudiantes algunos conflictos semióticos. La interpretación correcta de la definición viene dada en el apartado B, donde se resalta el carácter aleatorio de los extremos del intervalo y el porcentaje de intervalos, calculados a partir de diferentes muestras del mismo tamaño, que cubren el valor del parámetro. El distractor A evalúa el conflicto consistente en confundir la media muestral y poblacional (BEHAR, 2001), y es incorrecto porque la media muestral \bar{x} siempre cae dentro del intervalo, ya que está centrado en la misma. El distractor C considera el conflicto producido al suponer los extremos del intervalo constantes, identificado por López-Martín et al. (2019b), mientras que el apartado D desvela la interpretación bayesiana del intervalo (BEHAR, 2001; OLIVO y BATANERO, 2007).

En la Tabla 2 se presentan los resultados al primer ítem, cuya respuesta correcta fue elegida por más de la mitad de estudiantes de Bachillerato y el 70% de los de ingeniería. El error más frecuente fue dar una interpretación bayesiana al intervalo de confianza (distractor D), que pone de manifiesto la existencia de lo que Morey *et al.* (2016) denominan “falacia fundamental de la confianza” y que está ampliamente extendida entre los participantes de nuestro estudio. Olivo, Batanero y Díaz (2008) indican que no es posible calcular la probabilidad de que el intervalo de confianza contenga al parámetro, pues la confianza no está depositada en el intervalo, sino en el

método de construcción de los intervalos. Son pocos los estudiantes que responden utilizando los otros distractores.

Tabla 2: Resultados del ítem 1. Interpretación del intervalo de confianza

	Bachillerato ($n = 58$)		Ingeniería ($n = 37$)	
	Frecuencia	%	Frecuencia	%
A	1	1,7	2	5,7
B	32	55,2	25	71,0
C	2	3,4	1	2,9
D	22	37,9	14	40,0
En blanco	2	3,4	0	0,0

Fuente: Elaboración de los Autores

4.2 Resultados en el ítem 2

Ítem 2. Hemos calculado un intervalo de confianza al 90% utilizando el valor medio \bar{x} obtenido a partir de una muestra de 10 casos. Si incrementamos el tamaño de la muestra a 1000, y calculamos un segundo intervalo al 90% de confianza:

- (a) Debemos tener más confianza de que μ caerá en nuestro segundo intervalo.
- (b) Sabemos que el segundo intervalo será 10 veces más estrecho.**
- (c) Espero que ambos intervalos de confianza tengan la misma precisión.
- (d) El segundo intervalo de confianza es 10 veces más ancho que el primero.

Este ítem evalúa la comprensión de la relación entre el ancho del intervalo de confianza y el tamaño de la muestra. La solución correcta viene dada por el apartado B, ya que el intervalo se obtiene sumando y restando a la media muestral un valor que depende inversamente de la raíz cuadrada del tamaño muestral, por lo que un aumento de éste en dos órdenes de magnitud implica una disminución del intervalo a la décima parte. El distractor A, que evalúa la confusión entre confianza y amplitud, es incorrecto porque en el segundo intervalo el nivel de confianza no varía. Igualmente es incorrecto el apartado C, pues la precisión del intervalo sí viene influenciada por el tamaño de muestra. El apartado D muestra el error de suponer la relación opuesta entre tamaño de muestra y amplitud del intervalo.

Tabla 3: Resultados en el ítem 2 sobre tamaño de muestra y amplitud del intervalo

	Bachillerato ($n = 58$)		Ingeniería ($n = 37$)	
	Frecuencia	%	Frecuencia	%
A	19	32,8	13	37,1
B	15	25,9	9	25,7
C	10	17,2	7	20,0
D	12	20,7	9	25,7
En blanco	2	3,4	2	5,7

Fuente: Elaboración de los Autores

En la Tabla 3 se presentan los resultados de este ítem, cuya solución correcta fue elegida únicamente por una cuarta parte de los participantes en cada grupo, lo que pone de manifiesto una importante dificultad en la comprensión de que, al aumentar el tamaño de la muestra, fijando todos los demás factores, disminuye la anchura del intervalo. La respuesta mayoritaria fue la expresada por el distractor A, elegida por una tercera parte de los estudiantes e incluso un poco más en los de ingeniería. Por otro lado, el distractor D fue elegido por la quinta parte de estudiantes de Bachillerato y la cuarta de estudiantes de ingeniería. Finalmente, una sexta parte de los encuestados de Bachillerato y una quinta parte del alumnado de ingeniería ha respondido que el aumento de la muestra no conlleva una diferencia en la precisión del intervalo (distractor C), manifestando un desconocimiento, o bien sobre la noción de precisión del intervalo, o bien asociando el término “precisión” a los números empleados, la cual pudiera no cambiar en principio (YÁÑEZ y BEHAR, 2009).

Los porcentajes mostrados por los estudiantes universitarios son muy similares a los obtenidos entre los de Bachillerato, si bien es de destacar que en este último grupo se incrementan los porcentajes de error dados en todos los distractores. Es más, a diferencia del ítem anterior, aparecen respuestas en blanco, lo que demuestra que el alumnado presenta conflictos al interpretar cómo influye el tamaño muestral sobre la anchura del intervalo de confianza. Dichos conflictos podrían estar asociados con el hecho de que el alumnado comprende la conveniencia de tener muestras de tamaños mayores, pues éstas nos aportan más información que las muestras pequeñas. Podemos interpretar que se ha traducido esta propiedad a la idea de tener una mayor confianza en el intervalo, sin atender al hecho de que el nivel de confianza no ha variado, lo que ya fue comentado por Morey *et al.* (2016), quienes indicaron que no se comprende que la amplitud también depende del coeficiente de confianza.

4.3 Resultados en el ítem 3

Ítem 3. Si, manteniendo todos los demás datos fijos, el nivel de confianza se reduce (por ejemplo, del 90% al 80%):

- (a) El intervalo de confianza será menos preciso.
- (b) El intervalo de confianza será más ancho.
- (c) El intervalo de confianza será más estrecho.**
- (d) El intervalo de confianza será más preciso.**

Este ítem evalúa la comprensión de la relación entre el nivel de confianza y el ancho del intervalo. Cuanto mayor es la confianza, mayor es la anchura del intervalo, por lo que C y D son respuestas correctas y análogas, mientras que los distractores A y B muestran una relación entre confianza y precisión opuesta a la real.

En la Tabla 4 se presentan los resultados del tercer ítem, donde las opciones correctas han sido elegidas por más estudiantes de ingeniería, especialmente la C, siendo menos los estudiantes en los dos grupos que han elegido la opción D. Casi dos terceras partes del alumnado de Bachillerato y cuatro quintas partes del de ingeniería han asociado una reducción en el nivel de confianza con una reducción en la precisión, lo cual pone de manifiesto que los estudiantes asocian claramente estos conceptos — posiblemente se trate de un conflicto derivado del lenguaje, pues depositamos más confianza en aquellos procesos que anuncian mayor precisión en sus resultados, de acuerdo con Yáñez y Behar (2009). Si descartamos que una disminución en el nivel de confianza pueda traducirse en un intervalo de igual anchura o precisión, la respuesta A es contraria a C y la B opuesta a D. Es por ello curioso que los participantes no hayan detectado que esta cuestión debía tener dos respuestas correctas, lo cual se pone también de manifiesto al comparar los resultados tan similares de los apartados B y D, lo que debe estar relacionado con las dificultades de interpretación de la palabra “precisión”.

Tabla 4: Resultados obtenidos en el ítem 3 sobre nivel de confianza y amplitud del intervalo

	Bachillerato (n = 58)		Ingeniería (n = 37)	
	Frecuencia	%	Frecuencia	%
A	36	62,1	20	57,1
B	7	12,1	10	28,6
C	13	22,4	19	54,3
D	7	12,1	8	22,9
En blanco	1	1,7	0	0,0

Fuente: Elaboración de los Autores

Salvo en el primer descriptor, que fue seleccionado por alrededor del 60% en ambas muestras, los porcentajes observados entre una clase y otra de estudiantes son ciertamente diferentes. En general, el alumnado universitario prácticamente duplica los porcentajes de acierto frente al de Bachillerato. Interpretamos que esta diferencia podría ser debida, en parte, a una mayor competencia matemática entre el alumnado que eligió la modalidad de Ciencias en el Bachillerato en cuanto al manejo de propiedades básicas que son útiles a la hora de interpretar una expresión algebraica y cómo influyen sus diferentes factores. No obstante, el poco éxito del apartado D pone de manifiesto nuevamente la dificultad para asimilar con éxito la noción de “precisión” del intervalo de confianza, pudiendo existir un conflicto de carácter más lingüístico que matemático propiamente.

Entre los estudiantes de Bachillerato, sólo cuatro alumnos eligieron dos respuestas a la vez en este ítem y, en concreto, los cuatro se equivocaron eligiendo las opciones A y B, sin que ningún alumno escogiera simultáneamente las respuestas correctas C y D. Sin embargo, en el caso de la muestra universitaria sí hubo siete estudiantes que eligieron las dos respuestas correctas (C y D), pero también hubo ocho estudiantes (que representan un 22,9% de la muestra) que eligió las dos incorrectas (A y B).

4.4 Resultados en el ítem 4

Ítem 4. Se calculan intervalos de confianza del 95% con muestras de 100 elementos. ¿Cuál o cuáles de las siguientes afirmaciones son verdaderas?

- (a) Si la desviación estándar de la población disminuye, la anchura del intervalo de confianza no cambia.
- (b) Si la desviación estándar de la población disminuye, la anchura del intervalo de confianza disminuye.**
- (c) Si la desviación estándar de la población aumenta, la anchura del intervalo de confianza disminuye.
- (d) Si la desviación estándar de la población aumenta, la anchura del intervalo de confianza aumenta.**

Adaptado de Behar (2001), esta cuestión sirve para analizar la relación entre la anchura del intervalo y el valor de la varianza de la población y, al igual que el ítem anterior, tiene dos opciones correctas (B y D), cuyos enunciados dan a entender su equivalencia — si entendemos que la anchura del intervalo debe aumentar cuando aumenta la desviación estándar, entonces también debe disminuir la primera cuando disminuye la segunda. En este caso, la anchura del

intervalo de confianza sigue la expresión $2z_{\alpha/2} \sigma/\sqrt{n}$ y, por tanto, si la desviación estándar (σ) disminuye, la anchura del intervalo de confianza también disminuye. Los distractores A y C evalúan la falta de comprensión de esta relación entre desviación estándar y anchura del intervalo.

En la Tabla 5 se presentan los resultados al cuarto ítem, donde sorprende que haya alumnos que no han seleccionado simultáneamente las dos opciones correctas. Nuevamente, los porcentajes de acierto entre la muestra universitaria duplican a los obtenidos en la muestra de Bachillerato, posiblemente por su mejor preparación matemática.

Tabla 5: Resultados obtenidos en el ítem 4 sobre varianza de la población y anchura del intervalo

	Bachillerato ($n = 58$)		Ingeniería ($n = 37$)	
	Frecuencia	%	Frecuencia	%
A	3	5,2	1	2,9
B	14	24,1	17	48,6
C	27	46,6	10	28,6
D	17	29,3	20	57,1
En blanco	2	3,4	3	8,6

Fuente: Elaboración de los Autores

Aunque en los estudiantes de Bachillerato las frecuencias de los apartados B y D fueron similares, sólo hubo cinco que respondieron correctamente, es decir, B y D simultáneamente. La opción mayoritaria entre estos estudiantes fue el distractor C, el cual fue seleccionado por casi la mitad de los estudiantes (46,6%), no comprendiendo que la anchura del intervalo de confianza es directamente proporcional a la desviación estándar de la población. Este error no está descrito en investigaciones anteriores y supone una aportación original de nuestro trabajo.

Por el contrario, el alumnado universitario supo responder correctamente a este ítem en un porcentaje razonable, contando con 14 estudiantes que respondieron las dos opciones correctas a la vez, señal de que comprendían cómo influye la desviación estándar sobre la anchura del intervalo de confianza.

4.5 Resultados en el ítem 5

Ítem 5. Se ha tomado una muestra de tamaño $n = 100$ para calcular el intervalo de confianza, al 95%, para la velocidad media de los coches que pasan por una carretera. Se sabe que la desviación típica poblacional de estas velocidades es de 20 km/h.

- (a) El ancho del intervalo de confianza depende de la media muestral \bar{x} .
- (b) La media muestral \bar{x} es un extremo del intervalo de confianza.
- (c) **Cabe la posibilidad de que el verdadero valor de μ no esté contenido en el intervalo.**
- (d) **El ancho del intervalo no depende de la media muestral.**

Este ítem es de elaboración propia y evalúa la comprensión de que el ancho del intervalo no varía cuando se traslada la media puesto que está centrado en ella. Las respuestas correctas son tanto C como D. La primera porque, al ser aleatorios los extremos, algunos intervalos no cubren el valor del parámetro. Con ella pretendemos observar si el alumnado cree que el verdadero valor del parámetro poblacional siempre está contenido en el intervalo de confianza cuando el nivel de confianza es muy alto, error señalado previamente por Olivo (2008). La respuesta D también es correcta, puesto que la expresión de la anchura del intervalo de confianza, $z_{\alpha/2} \sigma/\sqrt{n}$, no depende de la media muestral. El distractor B es claramente falso, ya que, para obtener los extremos, siempre se suma y se resta una cantidad positiva a la media muestral, lo que obliga indiscutiblemente a que esta no pueda ser uno de los extremos. El distractor A es la negación del apartado D y, por tanto, es incorrecto.

El análisis de resultados del ítem 5 desveló que solo siete alumnos de Bachillerato — todos ellos del instituto B — eligieron conjuntamente las dos respuestas correctas, mientras que, del resto de aciertos, fueron más los que respondieron que el intervalo podría no contener a la media poblacional (36,2%) que los que indicaron que el ancho no depende de la posición de la media (29,3%). Respecto de los errores, la mitad pensó que el ancho del intervalo depende de la media muestral (distractor A), lo que significa que no comprenden la fórmula de cálculo. Este error no se ha descrito por otros investigadores y constituye una aportación de nuestro trabajo.

Por otro lado, el alumnado universitario muestra el conocimiento de que la media poblacional podría no pertenecer al intervalo de confianza (opción C) pero, por el contrario, son más los que consideran que el ancho del intervalo de confianza depende de la media muestral que los que no (51,4% vs. 42,9%; Tabla 6).

En términos generales, está claro que el alumnado comprende que el intervalo de confianza depende de la media muestral, que es una parte esencial de su expresión, pero globalmente desconoce que el ancho de dicho intervalo no depende de la misma. Igualmente indican no comprender la fórmula quienes eligieron el distractor B, pensando que la media está en un extremo — este porcentaje es significativo en la muestra de alumnado preuniversitario. Por

tanto, posiblemente no comprenden la fórmula ni el concepto de extremo del intervalo. Este es un tema en el que se debe ahondar más en clase, distinguiendo entre la expresión general del intervalo de confianza y el valor concreto de su anchura.

Tabla 6: Resultados obtenidos en el ítem 5 sobre la media muestral y la anchura del intervalo

	Bachillerato ($n = 58$)		Ingeniería ($n = 37$)	
	Frecuencia	%	Frecuencia	%
A	29	50,0	18	51,4
B	16	27,6	2	8,6
C	21	36,2	29	82,9
D	17	29,3	15	42,9
En blanco	1	1,7	0	0,0

Fuente: Elaboración de los Autores

4.6 Resultados en el ítem 6

Ítem 6. Se sabe que la altura media μ de los pinos de una zona forestal muy amplia es de 6.5 metros. ¿Cuál de las siguientes afirmaciones es verdadera?

- (a) Un posible intervalo de confianza para μ es [6.25, 8.4].
- (b) Un posible intervalo de confianza para μ es [6.75, 8.4].
- (c) Cualquier intervalo de confianza para μ que calculemos, asociado a una muestra aleatoria cualquiera, debe contener al valor 6.5.
- (d) En este contexto, no tiene sentido determinar el intervalo de confianza, ya que se conoce la altura media de todos los pinos.**

Esta cuestión también es de construcción propia y con ella pretendemos conocer si el alumnado comprende que el intervalo de confianza solo tiene sentido para estimar un valor desconocido de un parámetro, siendo, por tanto, D la respuesta correcta. Los distractores A y B son incorrectos porque no tiene sentido calcular un intervalo cuando se conoce la media de la población; además, la opción C es falsa porque, en ocasiones, el intervalo de confianza no contiene al verdadero valor del parámetro.

En la Tabla 7 se presentan los resultados al último ítem. La opción correcta (D) fue elegida por el 36,2% del alumnado de Bachillerato y el 40% de universitarios, resultado que nos parece muy pobre porque esta cuestión hace referencia directa a la justificación del estudio del tema: determinamos un intervalo de confianza cuando deseamos hacer una estimación de un valor

desconocido, de tal modo que si dicho valor es ya conocido, no tiene sentido llevar a cabo un proceso — en ocasiones, costoso y complicado — para su estimación.

Las opciones elegidas mayoritariamente por el alumnado en las dos muestras son diferentes pero incorrectas igualmente. Entre el alumnado preuniversitario, la opción mayoritaria fue el distractor C, hecho que posiblemente se deba a que los estudiantes centran su estudio en el proceso de cálculo del intervalo de confianza, sin prestar demasiada atención a su interpretación y al tipo de situación donde hay necesidad de calcularlo — estimación de un parámetro desconocido. El alumnado ha asimilado que la media muestral pertenece siempre al intervalo de confianza — de hecho, es su centro —, pero confunde media muestral (\bar{X}) y media poblacional (μ) a pesar de haberse explicitado esta última en el enunciado. Este error fue ya señalado tanto por Behar (2001) como por Olivo (2008). Olivo et al. (2008), donde los autores describen que un 25% de sus 252 estudiantes de ingeniería mostró confusión entre las ideas de estadístico y parámetro. Esta confusión no solo tiene lugar con la media, sino que en ese mismo estudio se pone de manifiesto que un 11.5% del alumnado confunde las desviaciones típicas de la población y de la distribución muestral.

Tabla 7: Resultados obtenidos en el ítem 6 sobre la necesidad de usar intervalos de confianza

	Bachillerato ($n = 58$)		Ingeniería ($n = 37$)	
	Frecuencia	%	Frecuencia	%
A	10	17,2	19	54,3
B	5	8,6	5	14,3
C	23	39,7	12	34,3
D	21	36,2	14	40,0
En blanco	1	1,7	0	0,0

Fuente: Elaboración de los Autores

Más de la mitad de los estudiantes universitarios eligieron el distractor A, que podría ser razonable si no conociésemos la media poblacional $\mu = 6,5$, pero en este caso el cálculo del intervalo carece completamente de sentido cuando, a priori, ya se conoce el valor del parámetro. Por esto mismo no habría razón para optar por el distractor B, que fue elegido por pocos estudiantes. Sin embargo, la principal razón por la que decidimos incluir este ítem en el cuestionario es saber si el alumnado llegaba a percibir la necesidad de usar los intervalos de confianza, y discriminar cuándo no es útil, hecho del que no tenemos constancia en investigaciones previas y que constituye otra aportación novedosa del presente trabajo.

4.7 Síntesis de resultados

Con objeto de comparar los resultados globalmente en los dos grupos se calificó cada una de las opciones correctas elegidas con 1 punto, obteniéndose una nueva variable “puntuación total” que puede variar teóricamente entre 0 y 9 puntos, puesto que hay tres ítems con dos opciones correctas, y se representó gráficamente la distribución de las puntuaciones en cada grupo con diagrama de barras (Gráfico 1) y de cajas (Gráfico 2). En lo Gráfico 2 se observa un mejor desempeño de los estudiantes de ingeniería, algunos de los cuáles alcanzan las puntuaciones 8 y 9, mientras que hay estudiantes de Bachillerato que no son capaces de identificar ninguna o sólo una respuesta correcta. Además, esta diferencia entre grupos también se evidencia al comparar la moda, pues en el de estudiantes de ingeniería es de 5 respuestas correctas (algo más de la mitad) mientras que la de los estudiantes de Bachillerato es únicamente de dos.

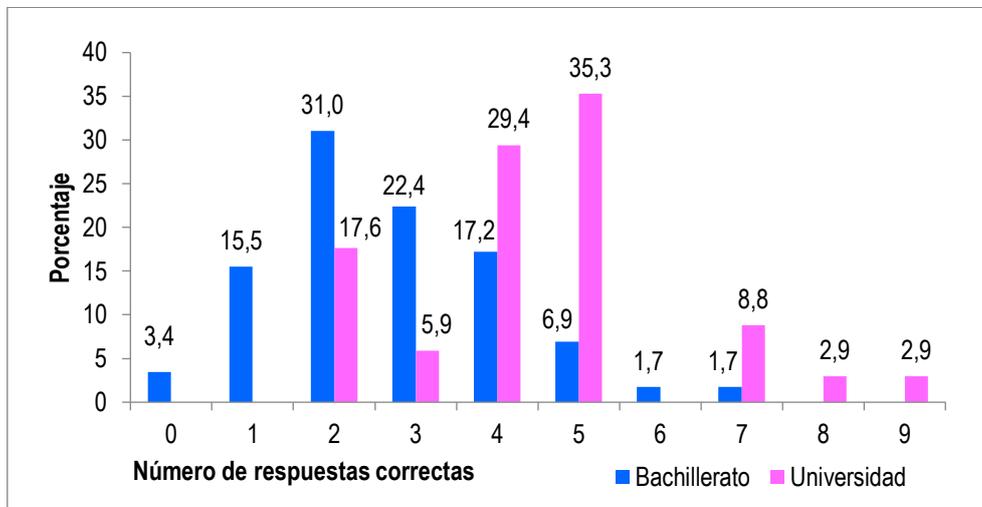


Gráfico 1: Distribución de la puntuación total por grupo (Elaboración de los Autores)

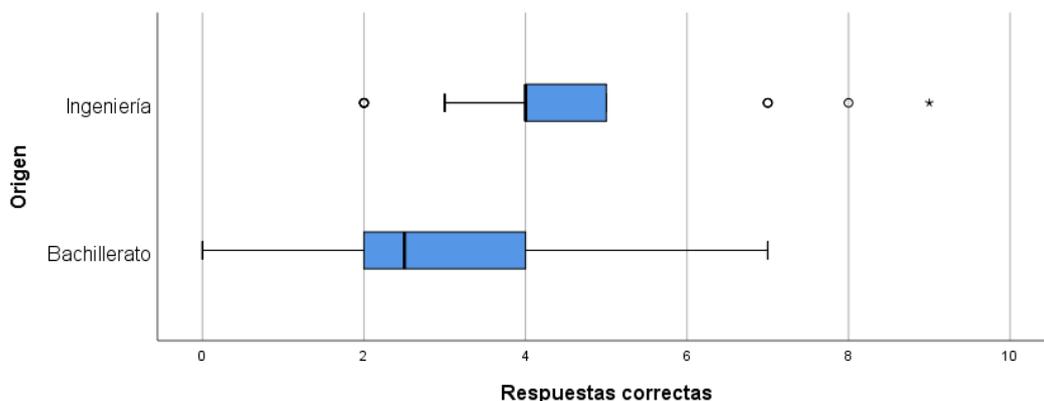


Gráfico 2. Distribución de puntuaciones en cada grupo (Elaboración de los Autores)

En lo Gráfico 2 se puede observar el valor superior de la mediana y los cuartiles en el grupo de ingeniería, donde el 50% de los estudiantes alcanza una puntuación de 4 a 5, mientras que el 50% de los estudiantes de Bachillerato alcanza un valor entre 2 y 4 respuestas correctas.

Estas diferencias visibles en los gráficos se pueden contrastar mediante el test t de diferencia de medias en muestras independientes, donde la media de los estudiantes de ingeniería es superior a las de Bachillerato (Tabla 8). En concreto se obtiene una diferencia de medias igual a 1,75 y un valor $t = 5,229$ que es altamente significativo ($p < 0,001$), siendo el intervalo de confianza para la diferencia de medias igual a (1,094- 2,406).

Tabla 8: Estadísticos resumen de las puntuaciones totales por grupo

Grupo	N	Media	Desviación típica	Error típico
Ingeniería	35	4,46	1,70	0,29
Bachillerato	58	2,71	1,44	0,19

Fuente: Elaboración de los Autores

Por otro lado, en la Tabla 9 se muestra el porcentaje de estudiantes en cada grupo que presentó los principales conflictos semióticos observados en el trabajo, ordenados descendientemente por frecuencia de aparición en los estudiantes de Bachillerato. Se indica también el ítem y distractor donde se detectó el conflicto.

Observamos un alto porcentaje de estudiantes que muestran conflictos ya detectados en investigaciones previas, como considerar que una reducción en el nivel de confianza implica una reducción en la precisión, confundiendo confianza y precisión, descrito en Yáñez y Behar (2009). Otro conflicto frecuente es considerar que el intervalo de confianza debe contener siempre a la media poblacional, lo que supone dar una interpretación determinista al mismo, dificultad que encontraron López-Martín *et al.* (2019a) en su trabajo, aunque en menor porcentaje al nuestro (11% vs. 40-34%). Igualmente, un alto porcentaje confunde confianza con probabilidad de que el parámetro caiga en el intervalo, dando una interpretación bayesiana al éste, error citado por Behar (2001), Morey *et al.* (2016) y Olivo (2008). Es decir, estos estudiantes interpretan el coeficiente de confianza como una probabilidad a posteriori que sólo puede ser calculada en el método bayesiano.

Tabla 9: Porcentajes de estudiantes que presentan diferentes conflictos semióticos

Conflicto semiótico	Ítem (distractor)	Bachillerato	Ingenieros
Reducir el nivel de confianza reduce la precisión	3 (A)	62,1	57,1
La amplitud depende de la media muestral	5 (A)	50,0	51,4
Disminuir la desviación típica aumenta el intervalo	4 (C)	46,6	28,6
El intervalo siempre contiene la media poblacional	6 (C)	39,7	34,3
Interpretación bayesiana del intervalo	1 (D)	37,9	40,0
Aumentar la muestra aumenta la amplitud del intervalo	2 (D)	37,9	25,7
Aumentar la muestra aumenta siempre la confianza	2 (A)	32,8	37,1
La media muestral es un extremo del intervalo	5 (B)	27,6	8,6
Calcular el intervalo cuando se conoce el parámetro	6 (A)	17,2	54,3
Ligar la precisión solo a la confianza	2 (C)	17,2	20,0
Reducir la confianza aumenta el intervalo	3 (B)	12,1	28,6
Disminuir la desviación típica no cambia el intervalo	4 (A)	5,2	2,8
Extremos del intervalo constantes	1 (C)	3,4	2,9
El intervalo contiene un % de veces la media muestral	1 (A)	1,7	5,7

Fuente: Elaboración de los Autores

Tienen variada frecuencia los conflictos ligados a la relación de la amplitud y precisión con diversos factores, tales como pensar que el ancho del intervalo aumenta cuando crece el tamaño de la muestra, error descrito en Fidler y Cumming (2005) entre el 20% de sus estudiantes de psicología (porcentaje muy parecido al nuestro). Igualmente, una parte de la muestra considera que el intervalo de confianza será más ancho cuando se reduce el nivel de confianza, descrito en Behar (2001) — en sentido contrario: cuando sube el nivel de confianza, baja la anchura del intervalo — y en Yáñez y Behar (2009), así como en Morey *et al.* (2016), Olivo (2008) y Olivo y Batanero (2007). Esta dificultad se observa también con menor frecuencia al ligar precisión y confianza o pensar que al cambiar la desviación típica no cambia el intervalo.

Al mismo tiempo no hay una gran proporción de estudiantes que consideren los extremos del intervalo constantes o que contenga la media muestral un porcentaje de veces, en contra de lo obtenido por Behar (2001) y Olivo (2008).

Además, hemos identificado nuevos conflictos semióticos no descritos anteriormente, algunos de los cuales se dan en alto porcentaje, como los siguientes:

- Considerar que la amplitud del intervalo depende de la media muestral, lo que podría venir explicado por el hecho de que la media muestral interviene en la fórmula de cálculo. Con menor proporción, algunos estudiantes piensan que la media muestral se sitúa en un extremo del intervalo de confianza.
- No comprender la finalidad de la estimación, aceptando un posible intervalo cuando se conoce el parámetro. Este error puede explicarse por una confusión entre estadístico y parámetro debido a que nos referimos a los dos con la misma palabra (media). La confusión de estadístico y parámetro fue señalada por Behar (2001), Harradine, Batanero y Rossman (2011), Olivo (2008) y Olivo *et al.* (2008), encontrando este último esa confusión en el 25% de sus estudiantes de ingeniería.
- Creer que el aumentar la desviación típica lleva a disminuir el intervalo, conflicto posiblemente ocurrido por recordar que en la fórmula de tipificación de la variable la varianza es un factor que divide.
- No relacionar las ideas equivalentes de “intervalo más estrecho” con “intervalo más preciso” — en el ítem 3, ningún alumno eligió ambas opciones correctas y equivalentes.
- Suponer que algunos intervalos no contienen a la media muestral — 1,7% en el distractor A del ítem 1 —, lo que demuestra que el alumnado no comprende que el centro del intervalo se sitúa en dicha media.
- Suponer que el tamaño de muestra aumenta la confianza.

Los conflictos semióticos anteriores aparecen con frecuencias muy similares en ambos grupos de estudiantes, a pesar de la mejor preparación matemática de los estudiantes de ingeniería. Son excepciones pensar que la media puede ser un extremo del intervalo (más frecuente en Bachillerato) o pensar que tiene sentido el intervalo cuando se conoce el parámetro (más frecuente en ingeniería).

5 Conclusiones

El estudio realizado informa de los diferentes conflictos semióticos que pueden tener los estudiantes al interpretar, en forma no acorde con la matemática, los elementos del intervalo de confianza y sus propiedades. Estos conflictos aparecen en dos grupos de estudiantes de diferente preparación, de lo que se deduce que hay aspectos que se deben mejorar en la enseñanza del

tema si se quiere conseguir unos profesionales estadísticamente bien formados.

Pensamos que la información descrita a lo largo del trabajo es útil para los profesores, algunos de los cuáles pudiera no ser consciente de la existencia de dichos conflictos en sus estudiantes. Más aún, la investigación didáctica muestra que estas interpretaciones erróneas de las expresiones y conceptos matemáticos también aparecen en ocasiones en los libros de texto que sirven de apoyo a la labor docente (GEA, LÓPEZ-MARTÍN y ROA, 2015).

Es importante profundizar en el tema, dada la dificultad en la comprensión del significado de este objeto matemático en toda su complejidad. Aunque el cálculo práctico es importante, no se deben obviar las propiedades que subyacen al intervalo de confianza y que le confieren su significado. En particular, es importante resaltar las siguientes:

- Un intervalo de confianza proporciona una respuesta al problema de estimación del parámetro poblacional en forma aproximada.
- Aunque la mayoría de intervalos de confianza generados con diferentes muestras aleatorias de la misma población contiene el verdadero valor del parámetro poblacional, algunos no lo contienen. El coeficiente de confianza nos da el porcentaje de intervalos que lo comprenden, pero no sabemos si el intervalo calculado en una determinada muestra lo contiene o no.
- Hay factores que influyen en el intervalo de confianza, tales como el coeficiente de confianza, tamaño de muestra y varianza de la población.
- En la enseñanza habitual se trabaja con un único intervalo de confianza asociado a una única muestra aleatoria y ello dificulta que el estudiante comprenda la interpretación frecuencial del intervalo de confianza y el carácter aleatorio de sus extremos. Por ello, proponemos el uso de hojas de cálculo, applets u otros recursos informáticos con los que el alumnado sea capaz de simular cientos de muestras aleatorias y, asociadas a cada una de ellas, un intervalo de confianza distinto para cada caso. Con ello podrán comprobar experimentalmente que el verdadero valor del parámetro no está siempre contenido en todos los intervalos simulados, ayudando a lograr una correcta interpretación del intervalo de confianza.
- Es importante también aclarar a los estudiantes diversos significados del término “precisión” y qué relaciones se establecen entre las nociones de “precisión de un

número” y “precisión del intervalo de confianza”.

Como conclusión general del estudio realizado destacamos la necesidad de dedicar más tiempo en clase a la interpretación del intervalo de confianza, pues será una noción de enorme importancia en los estudios posteriores del alumnado.

Agradecimientos

Proyectos TIN2017-89517-P y EDU2016-74848-P (AEI, FEDER) y Grupos de Investigación FQM-365 y FQM-126 (Junta de Andalucía).

Referencias

BEHAR, Roberto. *Aportaciones para la mejora del proceso de enseñanza-aprendizaje de la estadística*. 2001. 210f. Tesis (Doctoral en Estadística y Investigación Operativa) — Departament d'Estadística y Investigació Operativa. Universidad Politècnica de Catalunya. Barcelona.

CASTRO SOTOS, Ana Elisa; VANHOOF, Stijn; VAN DEN NORORGATE, Wim; ONGHENA, Patrick. [Student's misconceptions of statistical inference: A review of the empirical evidence from research on statistical education](#). *Educational Research Review*, v. 2, n. 2, p. 98-113, 2007.

CUMMING, Geoff. *Understanding the new statistics: effect sizes, Confidence intervals, and meta-analysis*. Nueva York: Routledge, 2012.

CUMMING, Geoff; WILLIAMS, Jennifer; FIDLER, Fiona. [Replication and researchers' understanding of confidence intervals and standard error bars](#). *Understanding Statistics*, n. 18, n. 3, p. 299-311, 2004.

FIDLER, Fiona; CUMMING, Geoff. Teaching confidence intervals: problems and potential solutions. In: 55th INTERNATIONAL STATISTICS INSTITUTE SESSION, 2005, Sidney. Proceedings of the 55th ISI. Sidney: International Statistical Institute, 2005.

GEA, María Magdalena; LÓPEZ-MARTÍN, María del Mar; ROA, Rafael. [Conflictos semióticos sobre la correlación y regresión en los libros de texto de Bachillerato](#). *Avances de Investigación en Educación Matemática*, v. 8, p. 29-49, 2015.

GODINO, Juan Díaz. Mathematical concepts, their meaning, and understanding. In: 20th CONFERENCE OF THE INTERNATIONAL GROUP FOR THE PSYCHOLOGY OF MATHEMATICS EDUCATION, 1996, Valencia. Proceedings of the 20th PME. Valencia: PME, 1996, p. 417-425.

GODINO, Juan Díaz. Un enfoque ontológico y semiótico de la cognición matemática. *Recherches en Didactique des Mathématiques*, v. 22, n. 2-3, p. 237-284, 2002.

HARRADINE, Anthony; BATANERO, Carmen; ROSSMAN, Allan. Students and teachers'

knowledge of sampling and inference. In: BATANERO, Carmen; BURRILL, Gail; READING, Chris (Ed.). *Teaching Statistics in School Mathematics-Challenges for Teaching and Teacher Education*, Springer: Netherlands, 2011, p. 235-246.

LÓPEZ-MARTÍN, María del Mar; BATANERO, Carmen; GEA, María Magdalena. Prospective high school teachers' interpretation of hypothesis tests and confidence intervals. In: JANKVIST, Uffe Thomas; VAN DEN HEUVEL-PANHUIZEN, Marja; VELDHUIS, Michiel (Ed.). *Proceedings of the Eleventh Congress of the European Society for Research in Mathematics Education*. Utrecht, the Netherlands: ERME. 2019a.

LÓPEZ-MARTÍN, María del Mar; BATANERO, Carmen; GEA, María Magdalena. [¿Conocen los futuros profesores los errores de sus estudiantes en inferencia?](#) *Bolema*, Rio Claro, v. 33, n. 64, p. 672-693, 2019b.

LÓPEZ-MARTÍN, María del Mar; BATANERO, Carmen; GEA, María Magdalena; ARTEAGA, Pedro. [Análisis de los problemas de inferencia propuestos en las Pruebas de Acceso a la Universidad en Andalucía](#), *Vidya*, Santa Maria, v. 36, n. 2, p. 409-428, 2016.

MAYO, Devorah G. [In defense of the Neyman-Pearson theory of statistics](#). *Philosophy of Science*, v. 48, n. 2, p. 269-280, jun. 1981.

MCED — MINISTERIO DE EDUCACIÓN, CULTURA Y DEPORTE, MECD. [Real Decreto 1105/2014, de 26 de diciembre, por el que se establece el currículo básico de la Educación Secundaria Obligatoria y del Bachillerato](#). Madrid: BOE, 3 ene. 2015.

MOREY, Richard D.; HOEKSTRA, Rink; ROUDER, Jeffrey N.; LEE, Michael D.; WAGENMAKERS, Eric-Jan. [The fallacy of placing confidence in confidence intervals](#). *Psychonomic Bulletin & Review*, v. 23, n. 1, p. 103-123, feb. 2016.

NEYMAN, Jerzy. [Outline of a theory of statistical estimation based on the classical theory of probability](#). *Philosophical Transaction of the Royal Society of London, series A, Mathematical and Physical Sciences*, v. 236, n. 767, p. 33-380, 1937.

OLIVO, Eusebio. [Significados del intervalo de confianza en la enseñanza de la ingeniería en México](#). 2008. 305f. Tesis (Doctoral en Ciencias de la Educación) — Departamento de Didáctica de la Matemática. Universidad de Granada. Granada.

OLIVO, Eusebio; BATANERO, Carmen. [Un estudio exploratorio de dificultades de comprensión del intervalo de confianza](#). *Unión*, v. 12, p. 37-51, dic. 2007.

OLIVO, Eusebio; BATANERO, Carmen; DÍAZ, Carmen. [Dificultades de comprensión del intervalo de confianza en estudiantes universitarios](#). *Educación Matemática*, v. 20, n. 3, p. 5-32, 2008.

THOMPSON, Bruce. [Effect sizes, confidence intervals, and confidence intervals for effect sizes](#). *Psychology in Schools*, v. 44, p. 423-432, apr. 2007.

YÁÑEZ, Gabriel; BEHAR, Roberto. Interpretaciones erradas del nivel de confianza en los intervalos de confianza y algunas explicaciones plausibles. In: GONZÁLEZ, María José; GONZÁLEZ, María Teresa; MURILLO, Jesús (Ed.). *Investigación en Educación Matemática: comunicaciones de los grupos de investigación. XIII Simposio de la SEIEM*, Santander, 2009.

YAREMKO, Robert M.; HARARI, Herbert; HARRISON, Robert C.; LYNN, Elisabeth. *Handbook of research and quantitative methods in psychology: for students and professionals*. Hilldale: Erlbaum, 2013.