

EL CICLO DE INVESTIGACIÓN ESTADÍSTICA EN EDUCACIÓN PRIMARIA: DECISIONES Y TENSIONES AL TRABAJAR CON UNA VARIABLE CUALITATIVA CONTEXTUALIZADA

The statistical investigation cycle in Primary Education: decisions and tensions when working with a contextualised categorical variable

Fraile-González, J.^a, Palop, B.^a, Garrido-Martos, R.^b. y Blázquez, R.^c

^aUniversidad Complutense de Madrid, ^bUniversidad Autónoma de Madrid, ^cCRA Sierra Oeste.

Resumen

El desarrollo de investigaciones estadísticas escolares en contextos socialmente relevantes ofrece oportunidades para trabajar prácticas estadísticas a partir de datos contextualizados. En este contexto, este estudio analiza las tensiones que emergen cuando alumnado de 6.º de Educación Primaria desarrolla una investigación estadística sobre una variable cualitativa en un contexto STEAM socialmente relevante. Se presenta un estudio de casos instrumental en el que se compara el proceso seguido por dos grupos de aula. Los resultados muestran que el desarrollo de la investigación se configura a partir de decisiones tempranas sobre la categorización y de la gestión de las tensiones derivadas, en parte, del fuerte anclaje contextual de la variable. A su vez, estas tensiones conducen a revisiones del propio ciclo y ponen de manifiesto tanto su carácter iterativo como el papel estructurante del contexto en investigaciones estadísticas escolares.

Palabras clave: educación estadística, investigación estadística, categorización de datos, GAISE, alfabetización estadística

Abstract

The development of school statistical research projects in socially relevant contexts offers opportunities to work with statistical practices based on contextualised data. In this context, this study analyses the tensions that emerge when 6th-grade students develop a statistical research project on a qualitative variable in a socially relevant STEAM context. An instrumental case study is presented, comparing the process followed by two classroom groups. The results show that the development of the research is shaped by early decisions regarding categorisation and the management of tensions generated in part by the variable's strong contextual anchoring. In turn, these tensions lead to revisions of the research cycle itself and highlight both its iterative nature and the structuring role of context in school statistical research.

Keywords: statistics education, statistical investigation, data categorization, GAISE, statistical literacy

INTRODUCCIÓN

La alfabetización estadística se ha consolidado como una competencia fundamental para la ciudadanía actual, ligada a la capacidad de interpretar y evaluar críticamente información basada en datos y de comunicar juicios fundamentados (Gal, 2002). Para avanzar hacia ella es preciso implicar al alumnado en investigaciones estadísticas desde la enseñanza básica (Bargagliotti et al., 2020), en las que el contexto ayuda a encontrar significado en los patrones observados e interpretar los datos en relación con la situación estudiada (Pfannkuch, 2011). Sin embargo, la estadística escolar tiende a reducirse a procedimientos descontextualizados y a actividades centradas en la representación de datos (Alsina et al., 2023), relegando procesos como la recogida, clasificación y organización de datos

(Alsina, 2025) y dificultando la articulación entre conocimiento estadístico y contextual necesaria para interpretar los datos en situaciones reales (Pfannkuch, 2011). En este sentido, la Educación STEAM, que promueve la resolución de problemas contextualizados y socialmente relevantes (Quigley et al., 2017), puede ofrecer un marco adecuado para desarrollar investigaciones estadísticas. No obstante, la relevancia del contexto puede influir en el desarrollo del ciclo de investigación estadística y generar tensiones entre el significado contextual y el tratamiento estadístico de los datos (Rodríguez-Muñiz et al., 2022). Resulta, por tanto, necesario profundizar en cómo se desarrollan investigaciones estadísticas en contextos socialmente relevantes (Burrill y Pfannkuch, 2024). El objetivo de este estudio es analizar las tensiones que emergen cuando alumnado de 6.º de Educación Primaria desarrolla una investigación estadística sobre una variable cualitativa en un contexto STEAM socialmente relevante.

MARCO TEÓRICO

Los problemas realistas que se abordan desde la Educación STEAM pueden ofrecer un marco para trabajar la estadística desde contextos relevantes (Berciano, 2025). De esta forma, se puede avanzar hacia una concepción de la estadística escolar como una actividad de investigación basada en datos (Alsina et al., 2023). El reporte GAISE II (Bargagliotti et al., 2020) establece un ciclo de resolución de problemas de cuatro etapas: formular preguntas estadísticas, recoger o considerar datos, analizar datos e interpretar resultados. Este ciclo se concibe como un proceso iterativo en el que las decisiones adoptadas en cada fase pueden requerir revisiones posteriores. El trabajo de Rodríguez-Muñiz et al. (2021) destaca la importancia de que los estudiantes participen en procesos de recogida, clasificación y organización de datos como parte del ciclo completo de investigación estadística. Estos procesos ofrecen al alumnado situaciones en las que la variabilidad requiere tomar decisiones para clasificar y organizar datos en contextos donde las categorías no son inmediatas ni inequívocas.

Dada la dificultad de aprender a categorizar los datos, que exige reconocer diferencias y similitudes, Rodríguez-Muñiz et al. (2021) proponen dos estrategias: clasificación supervisada (categorías previas y discusión de casos ambiguos) y clasificación no supervisada (categorías emergentes tras observar los datos). Los datos organizados mediante tablas de registro pueden facilitar la emergencia de tablas de frecuencias, en las que el recuento permite comprobar el número de valores observados en cada categoría (Alsina et al., 2023; Batanero, 2013). Estas tablas de frecuencias sustentan el análisis posterior mediante la creación de representaciones gráficas y la obtención de parámetros estadísticos (Bargagliotti et al., 2020). La interpretación de los resultados, en interacción con el contexto de los datos, puede revelar discrepancias, ambigüedades o insuficiencias en los datos recogidos. Esto puede conducir a profundizar en la indagación, buscar información adicional o reconsiderar las decisiones adoptadas (Frischemeier y Schnell, 2023).

Dado que los datos no son entidades abstractas sino representaciones de fenómenos contextualizados, su contexto constituye un componente estructurante del propio proceso investigativo (Cobb y Moore, 1997), ya que condiciona qué datos recoger, cómo analizarlos y cómo interpretar los resultados (Wild y Pfannkuch, 1999). En este sentido, Pfannkuch (2011) distingue entre el contexto del problema –del que proceden los datos– y el contexto experiencial del alumnado, señalando que la interacción entre ambos puede influir en el desarrollo del ciclo de investigación y en la construcción del razonamiento estadístico. Cuando la actividad estadística se sitúa en escenarios auténticos para trabajar con datos reales y próximos al alumnado, el contexto condiciona y puede enriquecer las decisiones adoptadas durante la investigación (Frischemeier y Schnell, 2023). El razonamiento estadístico exige articular simultáneamente una visión contextual y abstracta de los datos. No obstante, cuando este equilibrio no se mantiene, el foco puede desplazarse desde el razonamiento estadístico hacia el propio significado contextual de los datos (Rodríguez-Muñiz et al., 2022). Desde esta perspectiva, entendemos las discrepancias, ambigüedades y revisiones que emergen durante la investigación como tensiones propias del desarrollo del ciclo estadístico escolar, especialmente cuando el alumnado debe articular simultáneamente decisiones estadísticas y significados contextuales.

METODOLOGÍA

El estudio forma parte de una investigación basada en diseño (IBD) centrada en explorar cómo la Educación STEAM puede generar contextos significativos para desarrollar un ciclo de investigación estadística. Para ello, se diseñó un proyecto STEAM (sobre la identidad) articulado en torno a la pregunta reto *¿Quiénes somos los niños de la clase?*, que integra una tarea de investigación estadística estructurada según el ciclo del GAISE. Dentro de esta IBD se realizaron dos ciclos de implementación de la tarea. En esta comunicación se presenta un estudio de casos colectivo instrumental (Stake, 2006) orientado a analizar comparativamente el proceso de investigación de una misma variable cualitativa. El estudio se centra en dos grupos de 6.º de Educación Primaria de centros distintos (G1 y G2), correspondientes respectivamente al primer y segundo ciclo de implementación de la IBD. Ambos grupos analizaron la misma variable *Ocio y aficiones* en el contexto del proyecto. La implementación con G1 (25 estudiantes) se realizó en abril de 2025. En este grupo, la tarea se desarrolló principalmente en gran grupo con el acompañamiento del maestro. El alumnado únicamente había trabajado de forma superficial algunos parámetros estadísticos en cursos anteriores. A partir del análisis retrospectivo de este primer ciclo se introdujeron modificaciones en el diseño de la tarea dirigidas a favorecer una mayor autonomía del alumnado en la formulación de preguntas y en los procesos de recogida, organización y representación de datos, lo que implicó una ampliación de la temporalización de la tarea. Estas modificaciones dieron lugar a una segunda implementación, realizada con G2 (16 estudiantes) en febrero de 2026. En esta implementación, el alumnado trabajó en grupos de cuatro con mayor autonomía y el acompañamiento de la maestra. El grupo había trabajado previamente la creación y el análisis de gráficos estadísticos. La tarea tuvo una duración aproximada de dos horas en G1 y tres en G2.

Los datos proceden de grabaciones audiovisuales y producciones del alumnado. El análisis se desarrolló mediante la segmentación de las grabaciones en episodios significativos (Schoenfeld, 2000) en los que el alumnado toma o revisa decisiones vinculadas al proceso de investigación estadística. Los episodios se organizaron según las fases del ciclo GAISE y se analizaron narrativamente para reconstruir cómo se desarrollaba la investigación en cada grupo, atendiendo a las prácticas estadísticas emergentes (por ejemplo, formulación o revisión de preguntas, construcción de categorías o elección de instrumentos de recogida) y a las tensiones que surgían durante el proceso. Las producciones escritas asociadas a cada episodio se analizaron de forma integrada. Se realizó primero un análisis intracasos y, posteriormente, un análisis intercasos orientado a identificar patrones comunes y contrastes en las tensiones emergentes durante el desarrollo de la investigación estadística. Este análisis comparado permitió profundizar en la comprensión de dichas tensiones a partir de las convergencias y divergencias observadas entre ambos casos, así como orientar el rediseño de la tarea en los sucesivos ciclos de la IBD.

RESULTADOS

Formulación de preguntas

En G1 no se desarrolla esta fase, ya que la variable de estudio está previamente definida y la pregunta es planteada por el maestro. El trabajo se inicia directamente en la categorización de las aficiones propuestas por el alumnado.

Por otro lado, el proceso de G2 se configura como altamente negociado en la formulación de la pregunta. En un primer momento el alumnado explora distintas formulaciones posibles, como “¿tienes alguna afición?” o “¿qué haces por las tardes?”. La maestra guía al grupo, a través de preguntas, para avanzar hacia la formulación de una primera pregunta investigable: “¿cuáles son tus aficiones favoritas?”. A partir de esta formulación, el alumnado discute si la pregunta permitiría recoger la información que buscan y comienza a considerar la diversidad de respuestas posibles en la clase. En este proceso emerge una tensión en torno al significado de “aficiones favoritas” y a la posibilidad de reducir la respuesta a una única opción. Al debatir si formular la pregunta en singular

o en plural (“¿cuál es tu afición favorita?” o “¿cuáles son tus aficiones favoritas?”), una alumna argumenta que una persona puede tener varias aficiones, mientras que otra considera incoherente tener más de una favorita. Aunque el grupo se aproxima finalmente a una formulación en singular, esta discusión introduce la posibilidad de respuestas múltiples y anticipa tensiones que emergerán posteriormente durante la recogida de datos. La decisión sobre el número de respuestas posibles comienza así a configurarse como un elemento que condicionará el resto de la investigación.

Consideración y recogida de datos

En G1, la recogida se apoya en un dato previo de cada estudiante. Esta limitación a una única afición otorga estabilidad al posterior sistema de categorías. A partir de los datos individuales, el maestro dirige un proceso de categorización hasta obtener las categorías representativas de la clase mediante una agrupación guiada: distintas aficiones deportivas se unifican bajo la categoría “deportes” y cada estudiante sitúa su dato dentro de una de las categorías resultantes, cerrando progresivamente el sistema categorial. Esta estabilización temprana favorece una mayor coherencia técnica y fluidez en la recogida y representación posterior, aunque con un grado de autonomía limitado para el alumnado. La dinámica de recogida es dirigida por el maestro y consiste en que cada alumno sitúa un post-it junto a la categoría correspondiente. Durante la actividad, el alumnado demanda poder elegir más de una afición, por lo que el maestro permite dos; sin embargo, esta apertura a respuestas múltiples no va acompañada de una reapertura del sistema categorial, generando tensiones en fases posteriores.

En G2, el sistema de categorías se configura como permeable, revisable y construido desde una atención a la variabilidad. La preocupación por la cobertura del sistema se traduce en intentos de codificar respuestas concretas en categorías más generales, mediante una negociación continua para agrupar o añadir categorías que permitan representar a todo el grupo. Por ejemplo, se debate si “música va dentro de arte” o si “en deportes van todos los deportes”. Esta revisión incluye también la previsión de situaciones límite, incorporando opciones como “no tengo”, “no está ahí” o “prefiero no contestar”. Al planificar la recogida de datos emerge inicialmente la lista nominal, pero el grupo avanza hacia la tabla de doble entrada al considerarla un registro más eficaz. No obstante, la recogida se inicia sin haber estabilizado completamente el sistema categorial. Durante la recogida, aunque la pregunta se formula en singular (“afición”), varios estudiantes responden más de una opción y ellas lo registran, rompiendo la correspondencia entre alumno y registro, lo que compromete la organización posterior de los datos y afecta a la representación gráfica.

Análisis de datos

En G1, el análisis se apoya inicialmente en una lectura visual inmediata de la representación construida con *post-it*. Desde los primeros intercambios, varios alumnos identifican que “sale mucho deporte”, lo que muestra una rápida detección de la categoría más frecuente a partir de la disposición material de los datos. No obstante, la propia representación introduce una dificultad analítica: la longitud desigual de algunas filas, provocada por la separación entre post-it, genera una impresión visual que no siempre coincide con la frecuencia real. Esta tensión aparece cuando una categoría parece mayor que otra por su extensión en el espacio, hasta que un alumno corrige esa lectura afirmando que otra tiene más “porque los he contado”. De este modo, en G1 el análisis de los datos evoluciona de una apreciación visual global al recuento como criterio de validación. La representación facilita, por tanto, una entrada rápida al análisis, pero también obliga a contrastar la percepción inicial con el conteo efectivo de los datos.

Por su parte, en G2 el análisis se inicia sobre la tabla, antes de la construcción del gráfico. A partir de ella, el alumnado identifica la categoría mayoritaria (“deportes”), la minoritaria, los empates entre algunas categorías y la presencia de respuestas agrupadas en “no está ahí”. En la tabla aparecen además errores que el alumnado no detecta, como el registro de un estudiante que marca simultáneamente “no está ahí” y “música” (Figura 1). La transición de la tabla al gráfico surge a partir de la propuesta de la maestra de realizar una representación más visual; sin embargo, su construcción

genera varias dificultades. La recogida de más de una respuesta por alumno provoca problemas para determinar el total de referencia y establecer la escala del eje vertical. Además, se producen desajustes entre la tabla y la representación, ya que algunas categorías registradas—como “no tengo”, con frecuencia 1—no se incorporan de forma consistente al gráfico. En consecuencia, el análisis de los datos queda condicionado por las decisiones adoptadas en la fase previa de recogida y organización.

Figura 1
Registro conflictivo en la tabla de recogida de G2

[illegible]

Interpretación de resultados

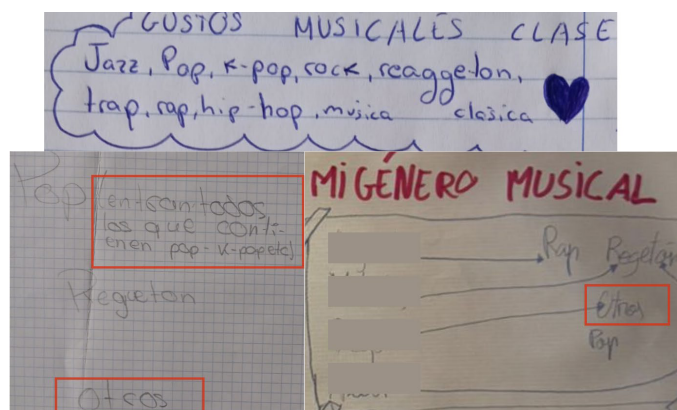
En G1, la interpretación está dirigida por el maestro. En primer lugar, se consolida “deportes” como la categoría más frecuente porque “La mayoría de la gente ha puesto post-it en deporte”, y el maestro introduce el término *moda*. Posteriormente, el foco se desplaza a la categoría “amigos”, cuya frecuencia se interpreta, según un alumno, como que “hay 10 personas que eligen amigos como ocio o afición favorita”. Esta lectura es cuestionada por otra alumna, que plantea que quizá “amigos” no sea una categoría equiparable a las demás, porque muchas actividades también se hacen con amigos. Finalmente, otro estudiante cuestiona la síntesis excesiva de la lista al afirmar que “hemos reducido mucho la lista”, y varios más señalan la ausencia de una categoría relativa a la música, mostrando que la categorización ha dejado fuera elementos importantes para representar al grupo.

Por su parte, G2 muestra una intensa problematización del proceso. En primer lugar, una alumna presenta la información que considera más relevante: “la mayoría ha puesto deporte”, y la maestra aprovecha esa lectura para introducir el término moda. A partir de ahí, se explicita qué significa que una categoría sea la que más se repite y el debate se desplaza hacia la categoría menos valorada. Cuando el grupo afirma que “ciencia” es la afición que menos personas realizan, se abre un debate con la maestra porque en la tabla había categorías previstas con frecuencia 0 que no habían representado. De este modo, emerge la distinción entre una categoría con menor frecuencia positiva y una categoría sin registros, formulada por la maestra como que no es “la que menos tiene”, sino “de las que tienen, es la que menos tiene”.

Como consecuencia de las tensiones identificadas durante la investigación, ambos grupos reabrieron el ciclo estadístico, aunque por motivos diferentes. En G1 se desarrolló en una sesión posterior una investigación centrada en la categoría no recogida, “música”, en la que el alumnado propuso, de forma autónoma, una pregunta—“¿Qué tipo de música te gusta más?”—, un sistema de categorías construido de forma autónoma que incorporaba el valor “otros” para evitar dejar respuestas fuera, como había ocurrido en la investigación anterior (Figura 2), un plan de recogida de datos basado en encuestas online, el uso de representaciones mediante gráficos y una interpretación preliminar a partir de resultados hipotéticos proporcionados por el maestro. Por su parte, en G2 se revisó la investigación en una sesión posterior, acordando recoger un único dato por alumno y realizando, en consecuencia, una nueva recogida, representación y análisis. En ambos casos, esta continuidad del trabajo muestra

una apropiación del ciclo estadístico por parte del alumnado y una devolución de la consigna, tanto en términos de completar la tarea como de dotarla de significado.

Figura 2
Sistema de categorías de la segunda investigación de G1



El análisis comparado muestra dos convergencias en el desarrollo de ambas investigaciones. En ambos grupos, la identificación de la categoría más frecuente emerge con rapidez a partir de la representación de los datos y es institucionalizada, según se había previsto en el diseño de la tarea de ambos ciclos. Asimismo, en ambos grupos la fase final no cierra la investigación, sino que deriva en nuevas investigaciones o revisiones del proceso. En G1 la nueva investigación sobre la música surge como respuesta a la limitación de los resultados para representar adecuadamente al grupo. Por su parte, en G2 emerge de forma no anticipada el deseo de rehacer la investigación cuando el alumnado detecta problemas en la recogida y representación de los datos. Esto muestra una apropiación del proceso de investigación presumiblemente favorecida por la relevancia contextual del proyecto.

Junto a estas convergencias, se observan diferencias –derivadas de decisiones de diseño de la tarea– en el desarrollo inicial del ciclo de investigación estadística, en el rol del docente y en la gestión del sistema de categorías. Tras la primera implementación se incorporó una fase inicial destinada a que el propio alumnado formulase la pregunta de investigación. Asimismo, mientras que en G1 se planificó una mediación docente más dirigida, en G2 el diseño se reorientó para promover una mayor autonomía del alumnado. Estas diferencias también se reflejan en la gestión del sistema de categorías. En G1, el maestro dirige su construcción y lo estabiliza tempranamente, mientras que en G2 la maestra promueve una dinámica más abierta, en la que las decisiones sobre la pregunta, las categorías y la recogida permanecen abiertas y revisables durante buena parte del proceso. Además, en G2 se anticipó la posibilidad de incorporar una categoría como “otros”, que ya emergió en la investigación de G1, apoyada en la idea de no dejar a ningún estudiante fuera del sistema categorial.

Además de las diferencias anticipadas, emergen dos cuestiones relevantes. Por un lado, en ambos casos se generaron, de forma no prevista, tensiones relacionadas con el sistema de categorías durante el desarrollo de la investigación. En G1, la tensión aparece durante la interpretación de los resultados, cuando el alumnado cuestiona si la categorización realizada representa adecuadamente al grupo debido a la ausencia de una categoría considerada relevante. En cambio, en G2 la tensión se hace visible durante el análisis y la representación, y deriva de no haber estabilizado suficientemente la pregunta y las categorías antes de la recogida de datos. Esto parece apuntar hacia un sistema categorial excesivamente rígido en G1 y hacia un inicio prematuro de la recogida de datos en G2.

Finalmente, el fuerte anclaje contextual de la variable, vinculada a la identidad y a las aficiones del propio grupo, desempeña un papel crucial en el desarrollo de la investigación. En G2, la importancia atribuida por el alumnado a sus aficiones dificulta estabilizar la pregunta y las categorías antes de la recogida de datos. Al mismo tiempo, en ambos casos el contexto favorece una implicación que va más allá de completar la tarea, llevando al alumnado a cuestionar las decisiones estadísticas adoptadas

–categorías, representaciones e interpretaciones– cuando considera que los resultados no representan adecuadamente al grupo. Como consecuencia, la investigación se amplía para abordar nuevas variables en G1 y se reinicia en G2 para rehacer todo el proceso.

DISCUSIÓN Y CONCLUSIONES

Los resultados muestran, en primer lugar, que las tensiones que han emergido durante la investigación estadística derivaban, por una parte, de decisiones adoptadas en torno a la categorización. En este sentido, la categorización estabilizada tempranamente ha facilitado la recogida y organización de los datos, pero ha generado tensiones posteriores relacionadas con la limitada capacidad del sistema categorial para representar la variabilidad del grupo. En cambio, la categorización más abierta y revisable ha favorecido la inclusión de esa variabilidad contextual, aunque ha generado tensiones en la representación y el análisis de los datos al no estabilizarse adecuadamente antes de la recogida. De este modo, los resultados refuerzan la idea de que la categorización no actúa únicamente como una forma de organizar la información, sino como un proceso de toma de decisiones que condiciona el desarrollo posterior de la investigación estadística (Alsina, 2025; Rodríguez-Muñiz et al., 2021).

Por otro lado, el contexto socialmente relevante de la investigación ha influido notablemente en el proceso. La importancia que las aficiones tenían para el alumnado generó tensiones entre el significado contextual de los datos y el razonamiento estadístico necesario para investigarlos. Sin embargo, este mismo contexto también favoreció una apropiación progresiva del proceso de investigación, llevando al alumnado a evaluar críticamente sus propias decisiones. De este modo, la relevancia del contexto puede favorecer el equilibrio entre razonamiento estadístico y conocimiento contextual (Rodríguez-Muñiz et al., 2022). Asimismo, la necesidad de responder al reto planteado en este contexto STEAM parece haber promovido la revisión y ampliación de las investigaciones (Frischemeier y Schnell, 2023). En este sentido, los nuevos ciclos de investigación constituyen una respuesta del alumnado ante la insuficiencia de los resultados para representar el contexto estudiado.

En conjunto, los resultados muestran el potencial de un contexto STEAM socialmente relevante para promover la implicación del alumnado y favorecer la emergencia de razonamientos estadísticos orientados a interpretar un fenómeno social próximo. Asimismo, la evolución entre ambos grupos se alinea con la intencionalidad del diseño de proporcionar una mayor autonomía al alumnado para desarrollar la investigación. En esta línea, un proceso de categorización autónomo y abierto puede promover el reconocimiento de la variabilidad, aunque requiere acompañamiento docente para estabilizar el sistema categorial antes de la recogida de datos. Finalmente, los resultados refuerzan la conveniencia, desde el punto de vista del aprendizaje, de permitir que el alumnado revise, amplíe o reinicie sus propias investigaciones.

Agradecimientos

Este estudio ha sido realizado en parte gracias al Proyecto de cooperación Erasmus KA2 de la Unión Europea (Project 2024-1-BE02-KA220-SCH-000245678), al Proyecto de Generación de Conocimiento 2024 (PID2024-155358NB-I00), a la ayuda FPU del Ministerio de Ciencia, Innovación y Universidades (FPU22/02803), al acuerdo de colaboración del Grupo de Investigación sobre Educación Científica y Matemática en la Sociedad (GIECMES) y al Colegio Rural Agrupado Sierra Oeste.

Referencias

Alsina, Á., Muñiz-Rodríguez, L., Rodríguez-Muñiz, L. J., García-Alonso, I., Vázquez, C. y López-Serentill, P. (2023). Alfabetizando estadísticamente a niños de 7-8 años a partir de contextos relevantes. *Revista Complutense de Educación*, 34(1), 95-108. <https://doi.org/10.5209/rced.77186>

- Alsina, Á. (2025). La investigación sobre estadística y probabilidad en edades tempranas: impacto en la escuela. En M. Burgos, M. C. Cañadas, J. M. Marbán e I. Polo-Blanco (Eds.), *Investigación en Educación Matemática XXVIII* (pp. 41–57). SEIEM.
- Bargagliotti, B., Franklin, C., Arnold, P., Gould, R., Johnson, R., Perez, L. y Spangler, D.A. (Eds.) (2020). *Pre-K–12 guidelines for assessment and instruction in statistics education II (GAISE II)*. American Statistical Association.
- Berciano, A. (2025). Retos de la investigación sobre sentido estocástico en proyectos STE(A)M de educación secundaria. En M. Burgos, M. C. Cañadas, J. M. Marbán e I. Polo-Blanco (Eds.), *Investigación en Educación Matemática XXVIII* (pp. 20–40). SEIEM.
- Burrill, G. y Pfannkuch, M. (2024). Emerging trends in statistics education. *ZDM–Mathematics Education*, 56, 19–29. <https://doi.org/10.1007/s11858-023-01501-7>
- Batanero, C. (2013). Sentido estadístico: componentes y desarrollo. *Actas de las Jornadas Virtuales en Didáctica de la Estadística, Probabilidad y Combinatoria*, 55–61. Universidad de Granada.
- Cobb, G. W. y Moore, D. S. (1997). Mathematics, Statistics, and Teaching. *The American Mathematical Monthly*, 104(9), 801–823. <https://doi.org/10.1080/00029890.1997.11990723>
- Frischemeier, D. y Schnell, S. (2023). Statistical investigations in primary school—the role of contextual expectations for data analysis. *Mathematics Education Research Journal*, 35(1), 217–242. <https://doi.org/10.1007/s13394-021-00396-5>
- Gal, I. (2002). Adults’ statistical literacy: Meanings, components, responsibilities. *International Statistical Review*, 70(1), 1–47. <https://doi.org/10.1111/j.1751-5823.2002.tb00336.x>
- Pfannkuch, M. (2011). The role of context in developing informal statistical inferential reasoning: A classroom study. *Mathematical Thinking and Learning*, 13(1–2), 27–46. <https://doi.org/10.1080/10986065.2011.538302>
- Quigley, C. F., Herro, D. y Jamil, F. M. (2017). Developing a conceptual model of STEAM teaching practices. *School Science and Mathematics*, 117(1–2), 1–12. <https://doi.org/10.1111/ssm.12201>
- Rodríguez-Muñiz, L.J., Muñiz-Rodríguez, L. y Aguilar-González, Á. (2021). El recuento y las representaciones manipulativas: los primeros pasos de la alfabetización estadística. *PNA*, 15(4), 311–338. <https://doi.org/10.30827/pna.v15i4.22511>
- Rodríguez-Muñiz, L. J., Muñiz-Rodríguez, L., García-Alonso, I., López-Serentill, P., Vázquez, C. y Alsina, Á. (2022). Navigating between abstraction and context in secondary school statistics education. *Culture and Education*, 34(3), 689–725. <https://doi.org/10.1080/11356405.2022.2058794>
- Schoenfeld, A. (2000). Models of the teaching process. *Journal of Mathematical Behavior*, 18(3), 243–261. [https://doi.org/10.1016/S0732-3123\(99\)00031-0](https://doi.org/10.1016/S0732-3123(99)00031-0)
- Stake, R. E. (2006). *Investigación con estudio de casos*. Ediciones Morata.
- Watson, J. M. (2006). *Statistical literacy at school: Growth and goals*. Lawrence Erlbaum.
- Wild, C. y Pfannkuch, M. (1999). Statistical thinking in empirical enquiry. *International Statistical Review*, 67(3), 223–248. <https://doi.org/10.1111/j.1751-5823.1999.tb00442.x>