



Mejora de los resultados educativos en contextos de exclusión social a través de un programa de educación STEM

Patricia Ayllón-Salas*

Departamento de Psicología Evolutiva y de la Educación de la Universidad de Granada

Mail: patriay@ugr.es

ORCID: <https://orcid.org/0000-0003-3986-225X>

Mirian Hervás-Torres

Departamento de Psicología Evolutiva y de la Educación de la Universidad de Granada

Mail: miriamhervas@ugr.es

ORCID: <http://orcid.org/0000-0002-8694-9393>

Javier Carrillo-Rosúa

Departamento de Didáctica de la Ciencias Experimentales de la Universidad de Granada

Mail: fjcarril@ugr.es

ORCID: <https://orcid.org/0000-0003-2889-3966>

Carmen Romero-López

Departamento de Didáctica de la Ciencias Experimentales de la Universidad de Granada

Mail: romero@ugr.es

ORCID: <https://orcid.org/0000-0003-4582-7397>

Juan Francisco Ruiz-Hidalgo

Departamento de Didáctica de las Matemáticas de la Universidad de Granada

Mail: jfruiz@ugr.es

ORCID: <https://orcid.org/0000-0002-4805-6922>

RESUMEN

En la actualidad, existe una gran necesidad de mejorar el nivel de cualificación en el ámbito científico-tecnológico. En este sentido, se desarrolló un programa de educación STEM para estimular el desarrollo científico-tecnológico en la educación obligatoria. Los objetivos de este estudio se dividieron en dos: (1) comprobar el impacto del programa para mejorar los resultados educativos del alumnado de educación primaria en riesgo de exclusión social, y (2) explorar los efectos del programa para incrementar el nivel de estrategias de aprendizaje autorregulado del alumnado universitario. Para el primer objetivo, la muestra estuvo compuesta por 50 estudiantes de primaria y, para el segundo objetivo, por 26 estudiantes del Grado en Educación Primaria. La intervención consistió en 20 sesiones de tutoría en las que el alumnado universitario realizaba actividades de refuerzo con el alumnado de educación primaria. Respecto a los resultados alcanzados con el alumnado de educación primaria, se encontraron diferencias a favor del grupo experimental en rendimiento escolar y actitudes hacia la ciencia. Sin embargo, los resultados con el alumnado universitario no fueron los esperados puesto que no se identificaron diferencias significativas. En general, los efectos del programa recomiendan continuar investigando para mejorar el programa y conseguir el impacto deseado.

Palabras clave: aprendizaje-servicio, educación STEM, formación inicial del profesorado, rendimiento escolar, aprendizaje-autorregulado.

*Autora de correspondencia: <https://orcid.org/0000-0003-3986-225X>

Recibido: 13/10/2024 – Aceptado: 23/03/2025

ISSN: 0210-2773

DOI: <https://doi.org/10.17811/rifie.21821>



Esta obra está bajo una licencia internacional Creative Commons
Atribución-NoComercial-SinDerivadas 4.0

Improving educational outcomes in contexts of social exclusion through a STEM education program

ABSTRACT

Currently, there is a clear necessity to enhance the qualifications of those working in the scientific-technological field. In this regard, a STEM education program was developed with the objective of stimulating scientific-technological development in compulsory education. This study had two objectives: (1) to assess the impact of the program on the educational outcomes of primary school students at risk of social exclusion, and (2) to examine the effects of the program on self-regulated learning strategies of university students. With regard to the first objective, the sample comprised 50 primary school students, while the second objective was addressed with a sample of 26 students enrolled in the Degree in Primary Education. The intervention comprised 20 tutoring sessions, during which university students conducted reinforcement activities with their primary education counterparts. With regard to the outcomes observed in the sample of primary school students, the experimental group demonstrated superior performance and more positive attitudes towards science compared to the control group. However, the results with university students did not align with expectations, as no significant differences were identified. In general, the effects of the program indicate a need for further research to improve the program and achieve the desired impact.

Keywords: service-learning, STEM education, initial teacher training, school performance, self-regulated learning.

1. Introducción

A lo largo de la última década la demanda de empleo en el sector científico-tecnológico se ha incrementado considerablemente (Organización para la Cooperación y el Desarrollo Económico, 2021), aunque la cantidad de estudiantes que accede y se gradúa en titulaciones universitarias y de formación profesional relacionadas con este sector está disminuyendo notablemente (Organización para la Cooperación y el Desarrollo Económico, 2024). Esta situación, resultado de la interacción de diversos factores de carácter social, cultural, económico y educativo (p. ej., el desinterés del alumnado hacia el campo científico-tecnológico durante la escolaridad obligatoria, las elevadas tasas de fracaso y abandono de estas titulaciones), está originando problemas de contratación y que determinados puestos no se estén cubriendo por la ausencia de profesionales con la cualificación requerida (National Science Board, 2022; Smith y White, 2019).

En este sentido, resulta más que evidente la necesidad de incrementar el número de alumnado y personal egresado en titulaciones científico-tecnológicas, además de mejorar su nivel de cualificación (Smith y White, 2019; Young *et al.*, 2017). Para ello, se han puesto en marcha diversas medidas y estrategias políticas, prácticas y de investigación que han dado lugar a diferentes iniciativas (Kearney, 2016; Smith y White, 2019; White y Smith, 2022; Young *et al.*, 2017) dirigidas a: reformar el currículo científico-tecnológico, sobre todo en materia de metodología docente; perfeccionar el desarrollo profesional de los docentes que imparten estas asignaturas; y mejorar la orientación vocacional y profesional que recibe el alumnado. Por ejemplo, en la Unión Europea se pueden enumerar bastantes iniciativas con estos propósitos (p. ej., Scientix o InGenious, lideradas por la European Schoolnet), mientras que en Estados Unidos se debe mencionar el plan «Educate to Innovate», así como los diferentes programas vinculados al National Science and Technology Council Committee on Science, Technology, Engineering, and Mathematics (STEM) Education (Caprile *et al.*, 2015; Kearney, 2016; Young *et al.*, 2017). No obstante, en España, de forma sorprendente, no se dispone de una estrategia específica para afrontar esta necesidad, si bien es cierto que la normativa educativa hace especial hincapié en la mejora de la enseñanza en el ámbito científico-tecnológico y cuenta con determinadas instituciones que están ofreciendo formación y entrenamiento a docentes (p. ej., Instituto Nacional de Tecnologías Educativas y de Formación del Profesorado) (Kearney, 2016), además del nuevo portal de la Alianza STEAM por el talento femenino.

La mayoría de estas iniciativas se han desarrollado en el ámbito educativo, desde la perspectiva de la educación STEM, concebida como un enfoque activo, interdisciplinar e integrador del proceso de enseñanza-aprendizaje de estas cuatro grandes áreas de conocimiento, fundamentado en el aprendizaje por indagación, que se orienta a la resolución de problemas e innovación con el propósito de incrementar el interés de los jóvenes por el ámbito científico-tecnológico (Jamali *et al.*, 2022; Roehrig *et al.*, 2021). Tal y como se refleja en la revisión Thibaut *et al.* (2018) donde recogen nueve categorías de prácticas instruccionales: integración de contenido STEM, enfoque en problemas, indagación, diseño, trabajo en equipo, centrado en el estudiante, de campo, evaluación y habilidades del siglo XXI, pone de manifiesto que el desarrollo del conocimiento parte de una construcción activa por los estudiantes, en la que sus ideas existentes y experiencias juegan un papel importante. Y es que los resultados de evaluación de los programas, proyectos y/o acciones de educación STEM implementados en las diferentes etapas educativas, tanto en tiempo formal como no formal e informal, revelan un impacto significativo en ciertas habilidades cognitivas (p. ej., atención) y no cognitivas (p. ej., autopercepciones, perseverancia, autocontrol, autorregulación, pensamiento crítico, resolución de problemas, etc.) así como en interés, conocimiento y rendimiento en asignaturas científico-tecnológicas (Eroğlu y Bektaş, 2022; Li *et al.*, 2022; Young *et al.*, 2017; Zhou *et al.*, 2024).

En este contexto, surge el programa de educación STEM que se presenta en este trabajo de investigación (<https://ntngranada.wordpress.com/>) un proyecto de colaboración interinstitucional entre las Universidades de Granada y Rutgers, en el que se ha pretendido examinar la aplicabilidad a otros contextos socioeducativos del modelo NtN-Rutgers (<http://ntn.rutgers.edu/>). El objetivo general del programa NtN-Rutgers es estimular el desarrollo del capital humano y social entre el alumnado de educación obligatoria en riesgo de exclusión social, integrando para ello actividades de refuerzo escolar con actividades científico-tecnoló-

gicas. Estas actividades se desarrollan en tiempo extraescolar y vacacional, siendo implementadas por el profesorado, personal y alumnado de Rutgers a través de instrucción directa, en el caso de las actividades de refuerzo escolar, y una metodología más experiencial para las actividades STEM, en las que se abordan situaciones de la vida cotidiana del alumnado con una exposición continua al método científico desde diferentes disciplinas. De hecho, los resultados de evaluación de las diferentes ediciones del NtN-Rutgers arrojan un alto nivel de eficacia, con un incremento considerable del conocimiento y el rendimiento escolar del alumnado en las áreas curriculares STEM, además de una mejora en sus habilidades cognitivas y no cognitivas (Camasso y Jagannathan, 2017; Jagannathan et al., 2018, 2019).

Obviamente, el programa NtN-Granada incorpora gran parte de los elementos y componentes del NtN-Rutgers que garantizan su éxito, pero de cara a incrementar su eficiencia, se determinó que fuera alumnado universitario, con un rol de tutores, el que implementara las actividades de refuerzo escolar y científico-tecnológicas con el alumnado de educación primaria, contribuyendo así en paralelo a la mejora de sus competencias profesionales. Para ello, se incorporaron al diseño de intervención del programa dos metodologías activas dirigidas a fomentar la adquisición de competencias, como son: el aprendizaje-servicio, que integra los elementos curriculares con el aprendizaje y la puesta en marcha de tareas útiles y solidarias en la comunidad (Chiva-Bartoll et al., 2020; Filges et al., 2022); y la tutoría entre iguales, que implica la creación de parejas de estudiantes, con una realización asimétrica, en la que el alumnado tutor es entrenado para facilitar ayuda y apoyo a otros estudiantes en un entorno planificado por profesionales (Arco et al., 2020; Thurston et al., 2021).

Las intervenciones que incorporan en su modelo lógico estas metodologías suelen implementarse con frecuencia en el ámbito educativo y social, principalmente por las evidencias disponibles sobre su eficacia para hacer frente a diferentes problemas (Fernández et al., 2018). Entre los múltiples beneficios de su uso se han identificado, entre otros, una mejora de los resultados educativos de los participantes, de sus habilidades y relaciones sociales, un aumento en indicadores de bienestar y competencias técnicas, así como mejoras en las habilidades profesionales y la satisfacción de los educadores implicados (Arco et al., 2020; Fernández et al., 2018; Filges et al., 2022; Foulkes y Naylor, 2022). También se dispone de evidencias sobre los beneficios que estas prácticas generan a nivel relacional en los centros que las desarrollan, mejorando también la estructura organizativa y participación de los miembros de las comunidades (Andrade et al., 2022; Fernández et al., 2018; García-López et al., 2023).

En definitiva, los objetivos de este estudio fueron, por un lado, determinar la efectividad del NtN-Granada para incrementar el interés y rendimiento escolar en las asignaturas STEM del alumnado de educación primaria, y, por otro lado, explorar sus efectos para mejorar la motivación y estrategias de aprendizaje del alumnado universitario. Por su parte, las hipótesis fueron: (H1) existirán diferencias estadísticamente significativas en rendimiento escolar de las asignaturas científico-tecnológicas a favor del grupo experimental de alumnado de educación primaria como resultado de su participación en el programa una vez finalizado el curso escolar; (H2) como consecuencia de su participación en el NtN-Granada, se apreciará un incremento estadísticamente significativo de las actitudes hacia la ciencia del alumnado de educación primaria del grupo experimental respecto al grupo control una vez finalizado el programa; (H3) en el caso del alumnado tutor, como resultado del programa, se observarán mejoras estadísticamente significativas en motivación y estrategias de aprendizaje en la fase postest respecto a la fase pretest; y (H4) como consecuencia del programa, el 90% del alumnado participante estará dispuesto a participar en próximas ediciones del NtN-Granada.

2. Método

2.1. Participantes

La muestra estuvo compuesta por 76 participantes, 50 estudiantes de educación primaria y 26 estudiantes universitarios. El alumnado de educación primaria se dividió en dos grupos, el grupo experimental, que se compuso de 25 estudiantes, 11 mujeres y 14 hombres, con una media de edad de 10,52 años (rango de entre 10 y 11 años), y una distribución por nivel educativo de 14 estudiantes de quinto curso y 11 de sexto curso. Por su parte, el grupo control estuvo formado por 25 estudiantes, con la misma distribución por nivel educativo y sexo, e idéntica media y rango de edad que el grupo experimental.

Por otro lado, los 26 estudiantes universitarios, que actuaron como tutores, cursaban el Grado en Educación Primaria y presentaban una media de edad de 19,58 años, con un rango de entre 17 y 39 años, mientras que su distribución por sexo era de 20 mujeres y seis hombres.

El procedimiento de selección de la muestra fue no probabilístico, de conveniencia (Kalton, 2020), e implicó cuatro etapas de selección. En la primera etapa, se seleccionaron dos centros educativos de educación primaria y el Grado en Educación Primaria atendiendo a los siguientes criterios: los limitados recursos financieros, materiales y humanos disponibles; el programa debía implementarse en una muestra de características educativas y sociodemográficas similares a la de los participantes de su homólogo americano (*i. e.*, quinto y sexto nivel en centros educativos ubicados en zonas urbanas en situación de pobreza o riesgo de exclusión), mientras que la titulación universitaria debía ser la que en mayor medida estaba ligada a la profesión docente en educación primaria; la naturaleza y objetivos del programa; y la disponibilidad y proximidad geográfica.

En la segunda etapa, una vez obtenidos los permisos éticos e institucionales oportunos, se llevó a cabo el plan de divulgación y captación, que consistió en: cuatro sesiones grupales de información con el alumnado de quinto y sexto curso de los centros educativos participantes, en las que se les informó de las características del programa a través de una breve actividad científica.

co-tecnológica, además de aportarles el Protocolo de consentimiento de familia (*i. e.*, documento similar a un contrato conductual en el que se especificaban los derechos y deberes de las familias, además de la autorización pertinente para que el centro educativo tramitara la inscripción de su hijo/a en el programa) (Fernández *et al.*, 2018); y dos sesiones grupales con el alumnado universitario, en las que se les informó de las características, condiciones y beneficios del programa, a la vez que se solicitó su participación voluntaria en el mismo.

En cuanto a la tercera etapa, 62 estudiantes de quinto y sexto curso fueron inscritos en el programa por los responsables de los centros educativos participantes, después de leer y aceptar las cláusulas del Protocolo de consentimiento del centro educativo (*i. e.*, documento similar a un contrato conductual en el que se concretaban el compromiso, derechos y deberes de los centros educativos) (Arco *et al.*, 2020; Fernández *et al.*, 2018), además de aportar el Protocolo de consentimiento de la familia, una copia del Boletín de notas del alumnado inscrito y el Protocolo de acceso (*i. e.*, autoinforme en el que los docentes aportaban la información demográfica, escolar y de interés del alumnado) (Arco *et al.*, 2020; Fernández *et al.*, 2018). Asimismo, 29 estudiantes universitarios se inscribieron voluntariamente en el programa después de leer y aceptar el Acuerdo para alumnado tutor (*i. e.*, documento similar a un contrato conductual en el que se especificaban sus derechos y deberes) (Arco *et al.*, 2020; Fernández *et al.*, 2018), aportar una copia de su expediente académico y completar el Protocolo para alumnado tutor que consistía en un autoinforme para obtener información demográfica, académica y de interés (Arco *et al.*, 2020; Fernández *et al.*, 2018), y el Cuestionario de Estrategias de Aprendizaje y Motivación (CEAM) (Roces *et al.*, 1995).

Por último, en la cuarta etapa, a partir de los 62 estudiantes de educación primaria inscritos en el programa, se crearon 26 pares asociados ($n = 52$) en base a dos grupos de variables control (Ato *et al.*, 2013) asociadas al rendimiento escolar (Choi y Calero, 2013; Fernández *et al.*, 2010): variables escolares: centro educativo, curso, grupo, Necesidades Específicas de Apoyo Educativo (NEAE), medidas de atención a la diversidad y rendimiento escolar previo; y variables demográficas: edad, sexo, nacionalidad, años de residencia en España e idioma empleado en casa. Los 10 estudiantes restantes fueron excluidos por no disponer de un par asociado apropiado, mientras que los miembros de uno de los pares asociados comunicaron su baja por incompatibilidad horaria.

Por otro lado, a partir de los 29 estudiantes universitarios inscritos en el programa, se seleccionó alumnado tutor ($n = 26$) bajo los siguientes criterios: presentar un rendimiento académico superior a los siete puntos, presentar puntuaciones centiles superiores a 50 en el CEAM (Roces *et al.*, 1995), demostrar interés, disponibilidad de tiempo, asistir a todas las sesiones y superar los diferentes ejercicios prácticos del curso de formación.

2.2. Materiales

La variable independiente de este estudio fue la participación del alumnado de educación primaria y alumnado universitario en el programa NtN-Granada.

En cuanto a las variables dependientes para el alumnado de educación primaria se trató de conocer (1) su rendimiento escolar entendido como el nivel de conocimientos que es demostrado en un área o materia cuando se compara con la norma de edad y nivel escolar (Jiménez, 2000), medido mediante las calificaciones extraídas del Boletín de notas; (2) las actitudes hacia la ciencia, entendidas como la relación con los comportamientos en el aprendizaje de las ciencias (matemáticas, ciencias naturales, etc.) y sus logros científicos a largo plazo (Alsop y Watts, 2003), la cual se medirá a través del Cuestionario de actitudes hacia la ciencia (Zhang y Campbell, 2011); y (3) satisfacción, entendida como la valoración realizada por el alumnado de educación primaria respecto a la adecuación de actividades realizadas, sus intereses personales, la implementación del programa, disposición a participar en futuras ocasiones y su satisfacción general, utilizando el Cuestionario de satisfacción.

En el caso del alumnado tutor, las variables dependientes del estudio fueron conocer (1) su motivación y estrategias de aprendizaje, definidas como el conjunto de aquellas estrategias que pueden permitir profundizar en el autoconocimiento de aspectos, por un lado, motivacionales y, por otro, de estrategias de estudio cuya finalidad será conocer cuáles son los procesos de regulación cognitiva y motivacional en beneficio de la calidad del aprendizaje (Pintrich *et al.*, 1991), a través del CEAM (Roces *et al.*, 1995); y (2) la satisfacción, considerada como la valoración del alumnado universitario respecto a la adecuación de las actividades, el desarrollo de competencias, sus intereses personales, la implementación del programa, su disposición a participar en futuras ocasiones y su satisfacción general, utilizando el Cuestionario de satisfacción.

Por lo tanto, los instrumentos utilizados en esta investigación fueron los siguientes:

- *Boletín de notas.* Copia original del boletín de notas del alumnado de educación primaria que se empleó para determinar su rendimiento escolar.
- *Cuestionario de actitudes hacia la ciencia* (Zhang y Campbell, 2011). Autoinforme formado por 28 ítems de cinco alternativas de respuesta agrupados en tres dimensiones (*i. e.*, sentimiento afectivo del alumnado en torno a la ciencia, tendencias de comportamiento del alumnado en el aprendizaje de la ciencia y juicio cognitivo del alumnado sobre la ciencia basado en sus valores y creencias sobre la ciencia). Este cuestionario se utilizó por ser uno de los pocos que mide las actitudes hacia la ciencia en el alumnado de educación primaria, además de presentar una fiabilidad (*i. e.*, alfa de Cronbach de entre 0,62 y 0,88 puntos en las diferentes escalas) y validez (*i. e.*, análisis factorial confirmatorio que justifica su estructura factorial) aceptables.

- CEAM (Roces *et al.*, 1995). Escala Likert constituida por 81 ítems de siete alternativas de respuesta agrupados en 15 escalas (*i. e.*, repaso, elaboración, organización, pensamiento crítico, autorregulación metacognitiva, manejo del tiempo y ambiente de estudio, regulación del esfuerzo, aprendizaje con iguales, búsqueda de ayuda, ansiedad, creencias de autoeficacia y de control del aprendizaje, metas de orientación intrínseca y extrínseca, y valor de las tareas), igual que la versión original (Pintrich *et al.*, 1993). Este cuestionario se empleó por ser uno de los pocos que mide motivación y estrategias de aprendizaje en el alumnado universitario, además de presentar una fiabilidad (*i. e.*, alfa de Cronbach de entre 0,48 y 0,86 para las escalas) y validez (*i. e.*, análisis factorial que demuestra la estructura factorial) aceptables.
- Cuestionario de satisfacción. Escala Likert compuesta de 18 ítems de estimación de uno a cinco puntos agrupados en dos áreas (*i. e.*, satisfacción con las actividades científico-tecnológicas y valoración global). Fue elaborado para recoger información sobre el nivel de satisfacción de los participantes, tomando como referencia los materiales de otros trabajos similares (Fernández *et al.*, 2018).

2.3. Diseño y procedimiento

Los diseños metodológicos adoptados en esta investigación fueron de carácter cuasiexperimental con grupo control no equivalente mejorado con técnicas de control estadístico para la H1, cuasiexperimental solamente postest para la H2, preexperimental pretest-postest para las H3 y preexperimental solamente postest para la H4 (Ato *et al.*, 2013).

Una vez ejecutado el muestreo, los miembros de los pares asociados de alumnado de educación primaria seleccionados fueron asignados aleatoriamente a la condición experimental y control a través de la aplicación <https://www.randomizer.org>. Posteriormente, se confirmó que ambos grupos eran equivalentes en torno a las variables control establecidas, ya que eran una constante análoga (*i. e.*, nacionalidad, años de residencia e idioma en casa) o mostraban la misma proporción (*i. e.*, centro educativo, curso, grupo, edad, sexo, presencia de NEAE, medidas de atención a la diversidad y asignaturas pendientes) en ambos grupos. Despues de emplear la prueba Kolmogorov-Smirnov y comprobar que la distribución de las puntuaciones era no normal, se emplearon análisis no paramétricos sobre rendimiento escolar previo en Matemáticas ($U = 264,50$, $p > 0,05$), Ciencias de la Naturaleza ($U = 246,50$, $p > 0,05$) y Ciencias Sociales ($U = 243,50$, $p > 0,05$). Los resultados no mostraron diferencias estadísticamente significativas para las variables estudiadas.

El plan de intervención se compuso de los contenidos, procedimientos y actividades extraídos de la bibliografía especializada (Andrade y Westover, 2020; Arco *et al.*, 2020; Eckert y Butler, 2021; Fernández *et al.*, 2018; Foulkes y Naylor, 2022; National Research Council, 2011). En este sentido, inicialmente se llevó a cabo el curso de formación para alumnado tutor, dirigido a potenciar en ellos aquellas competencias necesarias para desempeñar de la forma más eficaz y eficiente posible las funciones a desempeñar en el programa. Para ello, en cuatro sesiones de tres horas, a través de una metodología activa y participativa, se trabajaron los siguientes contenidos: sesión 1: presentación del equipo responsable del programa, participantes y del plan de formación, justificación del programa, y presentación y utilización del Cuaderno de trabajo (*i. e.*, material en el que se presentaba de forma estructurada cada una de las sesiones de tutoría) (Arco *et al.*, 2020; Fernández *et al.*, 2018); sesión 2: realización del análisis de necesidades del alumnado (*i. e.*, comentarios, análisis y revisión de agenda y material escolar para establecer y priorizar sobre deberes, dudas, problemas o dificultades sobre las diferentes asignaturas) y diseño y ejecución de actividades de refuerzo escolar; sesión 3: diseño y ejecución de actividades científico-tecnológicas; y sesión 4: evaluación y registro de las sesiones de tutoría.

Una vez seleccionado el alumnado tutor, se formaron 13 parejas, considerando su equivalencia en titulación, curso, grupo, disponibilidad horaria y afinidad académica y personal. Seguidamente, el alumnado del grupo experimental ($n = 25$) fue asignado a las parejas de alumnado tutor en función de su disponibilidad horaria (*i. e.*, a dos parejas se le asignaron tres, a ocho parejas dos y a tres parejas uno), para seguidamente implementar las sesiones de tutoría.

Las sesiones de tutoría se llevaron a cabo a lo largo del curso académico en tiempo extraescolar, con una frecuencia semanal en los espacios habilitados en cada centro educativo. En total, se ejecutaron 20 sesiones de tutoría, que se estructuraron en dos grandes bloques de actividades, con un pequeño receso entre ambos para que los participantes tomaran la merienda, fortalecieran sus vínculos y compromiso, y mejoraran así el clima de trabajo.

El primer bloque se destinó a actividades de refuerzo escolar (*i. e.*, realizar deberes, resolver dudas, problemas o dificultades, aumentar la práctica y el repaso en las diferentes áreas curriculares, especialmente instrumentales, mejorar la motivación, adaptar materiales, etc.), que se realizaron a través de instrucción directa (*i. e.*, establecimiento de condiciones óptimas para el desarrollo de la sesión y análisis de necesidades, objetivos y planificación de la sesión, aportación de información, ejemplificaciones y modelado, supervisión de la comprensión y práctica guiada, y práctica independiente). Por su parte, en el segundo bloque se ejecutaron las actividades científico-tecnológicas, ligadas a contenidos de Matemáticas, Ciencias de la Naturaleza y Ciencias Sociales (*i. e.*, digestión, aparato locomotor, anatomía de peces, materiales terrestres, células, función del microscopio, funcionamiento del microscopio, sistema circulatorio, biología vegetal, presión y atmósfera, microbios, artrópodos, gráficos y movimientos, gravedad, tipos de mezclas y su separación, densidad y flotabilidad, sistema respiratorio y salud, volumen y capacidad, sonido y luz, electricidad y magnetismo), mediante una metodología más experiencial, con exposición continua al método científico, como es el aprendizaje

por indagación (*i. e.*, identificación o establecimiento del problema/necesidad/objetivo, formulación de hipótesis, experimentación y recogida de información, discusión sobre los datos y establecimiento de conclusiones).

Las sesiones estaban estructuradas y secuenciadas en los Cuadernos de trabajo (*i. e.*, justificación curricular de las actividades científico-tecnológicas, los recursos materiales necesarios para su implementación y las actuaciones para su desarrollo, es decir, presentación e instrucciones iniciales, conceptos a abordar y secuencia de actuación), aunque el alumnado tutor tuvo que realizar adaptaciones a las características y necesidades de su alumnado, otorgándoles así más protagonismo y responsabilidad, como establece la literatura especializada (Arco *et al.*, 2020; Foulkes y Naylor, 2022). Asimismo, para facilitar su labor, disponían de una copia digital de los libros de texto de las diferentes asignaturas.

Por otro lado, en paralelo a la intervención, se desarrollaron diferentes medidas de seguimiento para vigilar las posibles desviaciones del programa: se incluyó un observador participante en la acción formativa; uno de los responsables del programa siempre estaba presente en las sesiones de tutoría; y cuatro sesiones grupales entre los responsables del programa y el alumnado tutor una vez realizadas la segunda, quinta, décima y vigésima sesión de tutoría, en las que se valoraron las actuaciones realizadas hasta ese momento y se aportaron las indicaciones y recomendaciones oportunas, además de establecer objetivos y tareas para próximas sesiones. Finalmente, con el plan de evaluación de resultados se tomaron medidas pretest-postest sobre las variables dependientes, para después comprobar la presencia o no de efectos significativos.

2.4. Análisis de datos

El análisis de datos conllevó las siguientes pruebas estadísticas para cada una de las hipótesis establecidas, después de emplear la prueba Kolmogorov-Smirnov para comprobar la normalidad en la distribución de las puntuaciones:

H1: prueba *U* de Mann-Whitney para la variable rendimiento escolar de las asignaturas científico-tecnológicas (*i. e.*, Ciencias de la Naturaleza, Ciencias Sociales y Matemáticas).

H2: prueba *t* para muestras independientes para la variable actitudes hacia la ciencia (*i. e.*, sentimiento afectivo, juicio cognitivo, y tendencias de comportamiento).

H3: prueba de rangos de Wilcoxon para muestras relacionadas para las variables de motivación y estrategias de aprendizaje (*i. e.*, orientación intrínseca, orientación extrínseca, valor de la tarea, control de las creencias de aprendizaje, autoeficacia, ansiedad, pensamiento crítico, autorregulación metacognitiva, tiempo y ambiente de estudio, estrategias de repaso, estrategias de elaboración, estrategias de organización, regulación del esfuerzo, aprendizaje entre iguales y búsqueda de ayuda).

H4: prueba de proporciones para la variable satisfacción (*i. e.*, satisfacción con las actividades científico-tecnológicas y valoración global).

El nivel de significación de las pruebas de comparaciones múltiples quedó ajustado a 0,01 para la H1 (0,05/5), 0,013 para la H2 (0,05/4) y 0,003 para la H3 (0,05/15).

Además, para las H1, H2 y H3 se calculó el tamaño del efecto determinado por el valor *d* de Cohen, mientras que la tasa de error por familia generada por las comparaciones múltiples fue controlada a través de la corrección de Bonferroni.

3. Resultados

A continuación, se presentan los resultados obtenidos según las variables indicadas anteriormente.

Los resultados sobre rendimiento escolar en las asignaturas STEM al finalizar el curso escolar del alumnado de educación primaria, H1, comprobada la ausencia de normalidad en la distribución de las puntuaciones, no revelan diferencias estadísticamente significativas entre las condiciones experimental y control (Tabla 1).

En cuanto a la H2, verificada la distribución normal de las puntuaciones, la prueba *t* para muestras independientes no revela diferencias estadísticamente significativas entre los grupos experimental y control del alumnado de educación primaria en las diferentes dimensiones y puntuación total del Cuestionario de actitudes hacia la ciencia (Zhang y Campbell, 2011) (Tabla 2).

Tabla 1
Comparaciones intergrupos sobre rendimiento escolar en asignaturas STEM

Asignatura / grupo	N	M	DT	U	p	d
Ciencias de la Naturaleza						
Experimental	25	7,40	1,68			
Control	25	6,96	1,74	267	0,37	0,26
Ciencias Sociales						
Experimental	25	7,20	1,73			
Control	25	6,80	1,85	277	0,48	0,22
Matemáticas						
Experimental	25	7,08	1,52			
Control	25	6,64	1,52	264	0,33	0,29

Nota. M = media; DT = desviación típica; U = prueba U de Mann-Whitney para dos muestras independientes; d = tamaño del efecto (valor d de Cohen's); p = significatividad estadística; *p < 0,01

Tabla 2
Comparaciones intergrupos sobre actitudes hacia la ciencia

Dimensión / grupo	N	M	DT	t	p	d
Sentimiento afectivo						
Experimental	25	48,33	4,87			
Control	25	45,95	5,65	1,6	0,12	0,45
Juicio cognitivo						
Experimental	25	24,88	3,23			
Control	25	23,88	3,39	1,07	0,29	0,30
Tendencias de comportamiento						
Experimental	25	38,76	6,39			
Control	25	39,33	6,40	-0,31	0,75	0,09
Total						
Experimental	25	111,98	12,61			
Control	25	109,16	13,36	0,77	0,45	0,22

Nota. M = media; DT = desviación típica; t = prueba t para muestras independientes; d = tamaño del efecto (valor d de Cohen); p = significatividad estadística; *p < 0,013

Respecto a la H3, las comparaciones pretest-postest sobre motivación y estrategias de aprendizaje del alumnado tutor, confirmada la ausencia de normalidad en la distribución de las puntuaciones, no arrojan resultados estadísticamente significativos (Tabla 3).

Por último, los resultados sobre satisfacción, H4, revelan que el alumnado del grupo experimental otorga al programa una valoración global media de 9,39 puntos ($DT = 0,86$), manifestando que volverían a participar en él un 89,3%, mientras que la satisfacción media general del alumnado tutor fue de 7,65 puntos ($DT = 2,02$), afirmando que, si se replicase el programa, volverían a participar en él un 65,38%. Los resultados de la prueba de proporciones no muestran diferencias estadísticamente significativas entre las proporciones esperadas y reales sobre la disposición a participar en próximas replicaciones del programa, tanto para el alumnado de educación primaria ($z = -3,00, p > 0,05$) como para el alumnado tutor ($z = -4,18, p > 0,05$).

Tabla 3
Comparaciones intragrupo sobre estrategias de aprendizaje y motivación

Escala / fase	N	M	DT	z	p	d
<i>Orientación intrínseca</i>						
Pretest	26	5,62	0,83	-		
Postest	26	5,81	0,91	-0,93	0,35	0,22
<i>Orientación extrínseca</i>						
Pretest	26	5,15	1,14	-		
Postest	26	4,91	1,20	-0,91	0,36	0,20
<i>Valor de la tarea</i>						
Pretest	26	5,68	0,85	-		
Postest	26	5,21	1,01	-2,85	0,00	0,50
<i>Control de las creencias de aprendizaje</i>						
Pretest	26	5,21	0,84	-		
Postest	26	5,28	0,94	-0,31	0,75	0,08
<i>Autoeficacia</i>						
Pretest	26	5,88	0,71	-		
Postest	26	5,79	0,68	-0,86	0,39	0,13
<i>Ansiedad</i>						
Pretest	26	3,05	1,30	-		
Postest	26	3,28	1,47	-1,07	0,29	0,16
<i>Pensamiento crítico</i>						
Pretest	26	4,93	1,24	-		
Postest	26	4,77	1,19	-0,69	0,49	0,13
<i>Autorregulación metacognitiva</i>						
Pretest	26	5,27	0,57	-		
Postest	26	5,26	0,54	-0,97	0,33	0,02
<i>Tiempo y ambiente de estudio</i>						
Pretest	26	4,88	0,38	-		
Postest	26	4,79	0,41	-0,89	0,37	0,23
<i>Estrategias de repaso</i>						
Pretest	26	5,07	1	-		
Postest	26	5,15	0,96	-0,15	0,88	0,08
<i>Estrategias de elaboración</i>						
Pretest	26	5,75	0,95	-		
Postest	26	5,84	0,84	-0,35	0,73	0,10
<i>Estrategias de organización</i>						
Pretest	26	5,5	0,88	-		
Postest	26	5,42	0,95	-0,64	0,52	0,09
<i>Regulación del esfuerzo</i>						
Pretest	26	3,93	0,45	-		
Postest	26	4,17	0,57	-2,38	0,01	0,47
<i>Aprendizaje entre iguales</i>						
Pretest	26	4,97	0,25	-		
Postest	26	5,01	0,21	-0,44	0,66	0,17
<i>Búsqueda de ayuda</i>						
Pretest	26	5,41	0,57	-		
Postest	26	5,32	0,68	-0,10	0,92	0,14

Nota. M = media; DT = desviación típica; z = prueba z de Wilcoxon; d = tamaño del efecto (valor d de Cohen); p = significatividad estadística; *p < 0,003

4. Discusión

Este programa de educación STEM se diseñó como una experiencia educativa innovadora internivelar dirigida a mejorar el interés y el rendimiento escolar del alumnado de educación primaria en las áreas curriculares STEM, así como a ofrecer al alumnado de titulaciones universitarias relacionadas con la profesión docente la posibilidad de participar en actividades de aprendizaje de carácter co-curricular, que les permita mejorar sus competencias académicas. Por supuesto, todo ello en colaboración con los responsables y docentes de los centros educativos participantes, y bajo la supervisión en todo momento del profesorado universitario responsable del programa. En este sentido, a partir de los resultados obtenidos, se pueden establecer las siguientes conclusiones: puesto que no se aprecian diferencias estadísticamente significativas entre los grupos experimental y control en cuanto a su rendimiento escolar en asignaturas STEM y actitudes hacia la ciencia una vez finalizado el programa, las H1 y H2 se rechazan; al no apreciarse diferencias pretest-postest estadísticamente significativas entre las puntuaciones directas logradas por el alumnado tutor en motivación y estrategias de aprendizaje, la H3 no se acepta; y debido a que el tanto por ciento de alumnado que está dispuesto a participar en próximas ediciones del programa es inferior al 90%, la H4 también se rechaza.

No obstante, si para el contraste de hipótesis se emplean pruebas diferentes a las tradicionales de significación estadística, tal y como recomienda la literatura especializada (Coe, 2002), la conclusión en torno a las H1 y H2 es muy diferente. La interpretación de los valores del tamaño del efecto conseguidos para las variables dependientes de ambas hipótesis revela que un miembro hipotético del grupo experimental puede alcanzar un rendimiento escolar en las asignaturas STEM superior al 58% del alumnado del grupo control, con porcentajes que pueden llegar incluso a ser superiores al 66% en el caso de actitudes hacia la ciencia (p. ej., sentimiento afectivo). De hecho, aunque el tamaño del efecto que se ha logrado en estas variables debe interpretarse como pequeño (Cohen, 1988) o pequeño-mediano (Sawilowsky, 2009), Coe (2002) establece que un cambio superior a 0,10 puntos puede considerarse como una mejora muy significativa, más aún si esa mejora es uniforme a todos los miembros del grupo y se mantiene en el tiempo. Además, la literatura indica que los tamaños de efecto alrededor de 0,2 puntos tienen una mayor importancia práctica o clínica cuando se basan en medidas de resultados que son más difíciles de cambiar, como es el caso del rendimiento escolar (Durlak, 2009). Por tanto, los resultados obtenidos parecen confirmar las previsiones planteadas inicialmente sobre los efectos de mejora que la participación en este tipo de programas conlleva para variables como el interés, conocimiento y rendimiento escolar del alumnado en asignaturas STEM (Camasso y Jagannathan, 2017; Chang *et al.*, 2022; Jagannathan *et al.*, 2018, 2019; Young *et al.*, 2017).

Ahora bien, la ausencia de efectos estadísticamente significativos en interés y rendimiento escolar en las asignaturas STEM pueden estar relacionados con la exposición al tratamiento por parte de los participantes, como señalan Camasso y Jagannathan (2017). Por ejemplo, las sesiones de tutoría comenzaron dos meses después del comienzo del curso escolar, principalmente por el elevado coste en tiempo y recursos humanos que tuvo la ejecución del muestreo y la configuración de los grupos. Este retraso generó que el alumnado de educación primaria del grupo experimental tuviera una exposición muy limitada a las actividades del programa antes de la evaluación del primer trimestre del curso escolar, con sus correspondientes consecuencias para el rendimiento escolar en esa primera evaluación. Asimismo, es necesario mencionar que el comportamiento de cierto alumnado (p. ej., impuntualidad, asistencia a las sesiones sin los materiales escolares necesarios, dificultades en la ejecución de las sesiones de tutoría, a veces por la falta de preparación previa) ha podido afectar a los resultados, así como la adecuación y ajuste de algunas actividades efectuadas en las sesiones de tutoría (p. ej., tipo, secuenciación y distribución de las actividades, con ciertas discrepancias con las actividades que se desarrollaban en el aula ordinaria, ante la ausencia de información y materiales que se empleaban en tiempo escolar) y los espacios en los que se desarrollaban.

En cuanto al comportamiento de las variables dependientes de la H3, tanto si se considera la significatividad estadística como el tamaño del efecto, la participación en el programa no ha tenido los efectos esperados sobre el desarrollo de competencias académicas del alumnado tutor. Los resultados obtenidos no coinciden con las evidencias empíricas que arroja la literatura especializada sobre los beneficios de la participación del alumnado en experiencias que incorporan en su modelo lógico el aprendizaje-servicio y/o la tutoría entre iguales (Andrade *et al.*, 2022; Arco *et al.*, 2020; Fernández *et al.*, 2018; Filges *et al.*, 2022). Estas discrepancias quizás se puedan explicar por la presencia de ciertas limitaciones metodológicas, principalmente ligadas al diseño adoptado y a los efectos de la administración de los instrumentos o sesgo de cambio de respuesta (*i. e.*, cuando la métrica de los participantes para responder preguntas entre las fases pretest y postest cambia debido a una mayor comprensión del concepto que se está investigando) (Fernández *et al.*, 2018). En este sentido, el alumnado tutor presentaba una percepción demasiado elevada sobre su motivación y estrategias de aprendizaje en la fase pretest, pero las experiencias quasi-profesionales a las que se enfrentaron, parecen haber provocado que se cuestionen su nivel de competencias académicas en la fase postest.

Por otro lado, la participación en el programa parece haber generado bastante satisfacción entre los participantes, especialmente entre el alumnado de educación primaria, que valoraron globalmente el programa como sobresaliente. Sin embargo, puede sorprender el elevado porcentaje de alumnado tutor que declinó participar en próximas ediciones del programa (*i. e.*, 34,62%), con argumentos ligados a la excesiva cantidad de tiempo, y a la implicación y esfuerzo que exigió su participación en el mismo.

Por supuesto, los hallazgos sobre la efectividad del programa deben interpretarse con precaución (*i. e.*, efectividad potencial), considerando las principales limitaciones del estudio. En esta línea, aunque el muestreo no probabilístico presenta muchas limitaciones, es útil, especialmente cuando los recursos, el tiempo y los recursos humanos son limitados (Etikan *et al.*, 2016). Por su parte, con relación al tamaño reducido de la muestra, se aplicaron reglas generales a la hora de calcular el tamaño de la muestra (*i. e.*, 20-40 participantes por grupo) (Whitehead *et al.*, 2016).

En definitiva, los resultados obtenidos con este trabajo de investigación pueden considerarse satisfactorios, pues han contribuido a determinar la viabilidad de los componentes del programa (p. ej., protocolo, muestreo, instrumentos de medida, recogida y análisis de datos), identificando sus debilidades y mejoras potenciales para el diseño y planificación de próximas replicaciones a mayor escala. Además, la participación en este tipo de experiencias parece particularmente relevante y útil para el alumnado universitario.

Estos resultados podrían mejorarse considerablemente con mayores niveles de participación, interés y colaboración entre los diferentes agentes educativos que han participado en las distintas fases del programa, como también señalan otros estudios del programa NtN-Rutgers (Jagannathan *et al.*, 2018), así como un mayor apoyo sistémico (Shah *et al.*, 2018). La asignación de recursos económicos dirigidos a fortalecer las funciones de gestión y rendimiento antes, durante y después de la implementación del programa beneficiará la fidelidad del programa, así como el impacto del programa en las variables dependientes. Esos recursos financieros también ayudarían a extender el programa más allá del curso escolar (p. ej., campamento de verano), en los momentos en los que el alumnado de educación primaria está libre de responsabilidades escolares y dispone de más tiempo para aprender y participar en propuestas educativas alternativas.

Financiación

Esta investigación se ha realizado dentro del proyecto 2015/22472 subvencionado por Rutgers, The State University of New Jersey (USA).

Referencias

- Alsop, S., y Watts, M. (2003). Science education and affect. *International Journal of Science Education*, 25(9), 1043-1047.
- Andrade, M. S., y Westover, J. H. (2020). Engaging Millennial Students through Community-Engaged Experiential Learning. *Experiential Learning & Teaching in Higher Education*, 2(2), 21-44
- Andrade, M., Workman, L., y Westover, J. H. (2022). Current practices for community-based learning in schools of business. *The International Journal of Management Education*, 20, 100597. <https://doi.org/10.1016/j.ijme.2021.100597>
- Arco, J. L., Fernández, F. D., y Hervás, M. (2020). Evidence-based peer-tutoring program to improve students' performance at the university. *Studies in Higher Education*, 45(11), 2190-2202. <https://doi.org/10.1080/03075079.2019.1597038>
- Ato, M., López, J. J., y Benavente, A. (2013). Un sistema de clasificación de los diseños de investigación en psicología. *Anales de Psicología*, 29(3), 1038-1059. <https://doi.org/10.6018/analesps.29.3.178511>
- Camasso, M. J., y Jagannathan, R. (2017). Improving academic outcomes in poor urban schools through nature-based learning. *Cambridge Journal of Education*, 48(2), 263-277. <https://doi.org/10.1080/0305764X.2017.1324020>
- Caprile, M., Palmén, R., Sanz, P., y Dente, G. (2015). Encouraging STEM studies in the labour market. European Parliament. <https://doi.org/10.2861/519030>
- Chang, S.-H., Yang, L.-J., Chen, C.-H., Shih, C.-C., Shu, Y., y Chen, Y.-T. (2022). STEM education in academic achievement: a meta-analysis of its moderating effects. *Interactive Learning Environments*, 32(6), 2401-2423. <https://doi.org/10.1080/10494820.2022.2147956>
- Chiva-Bartoll, O., Moliner, M. L., y Salvador-García, C. (2020). Can service-learning promote social well-being in primary education students? A mixed method approach. *Children and Youth Services Review*, 111, 104841. <https://doi.org/10.1016/j.childyouth.2020.104841>
- Choi, A., y Calero, J. (2013). Determinantes del riesgo de fracaso escolar en España en PISA-2009 y propuestas de reforma. *Revista de Educación*, 362, 562-593. <https://doi.org/10.4438/1988-592X-RE-2013-362-242>
- Coe, R. (2002). It's the effect size, stupid. What effect size is and why it is important. *Annual Conference of the British Educational Research Association, September 12–14*. University of Exeter.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2^a ed.). Lawrence Erlbaum Associates.
- Durlak, J. A. (2009). How to select, calculate and interpret effect sizes. *Journal of Pediatric Psychology*, 34(9), 917-928. <https://doi.org/10.1093/jpepsy/jsp004>
- Eckert, J., y Butler, J. (2021). Teaching and Leading for Exemplary STEM Learning. A Multiple-Case Study. *The Elementary School Journal*, 121(4), 674-699. <https://doi.org/10.1086/713976>
- Eroğlu, S., y Bektaş, O. (2022). The effect of 5E-based STEM education on academic achievement, scientific creativity, and views on the nature of science. *Learning and Individual Differences*, 98, 102181. <https://doi.org/10.1016/j.lindif.2022.102181>
- Etikan, I., Musa, S. A., y Alkassim, R. S. (2016). Comparison of convenience sampling and purposive sampling. *American Journal of Theoretical and Applied Statistics*, 5(1), 1-4. <https://doi.org/10.11648/j.ajtas.20160501.11>
- Fernández, F. D., Arco, J. L., Hervás, M., y Delgado, L. C. (2018). Transformar la educación superior y obligatoria a través de aprendizaje-servicio y mentoría. *Universitas Psychologica*, 17(4), 1-12. <https://doi.org/10.11144/Javeriana.upsy17-4.teso>
- Fernández, M., Mena, L., y Riviere, J. (2010). *Fracaso y abandono escolar en España*. Fundación La Caixa.
- Filges, T., Dietrichson, J., Viinholt, B. C. A., y Dalgaard, N. T. (2022). Service learning for improving academic success in students in grade K to 12: A systematic review. *Campbell Systematic Reviews*, 18, e1210. <https://doi.org/10.1002/cl2.1210>
- Foulkes, D., y Naylor, S. (2022). Exploring peer tutoring from the peer tutor's perspective. *Radiography*, 28(3), 793-797. <https://doi.org/10.1016/j.radi.2022.02.007>

- García-López, L. M., Gutiérrez, D., González-Martí, I., Segovia, Y., y MacPhail, A. (2023). Enhancing sport education experiences through servicelearning. *European Physical Education Review*, 29(2), 215-232. <https://doi.org/10.1177/1356336X221132770>
- Jagannathan, R., Camasso, M. J., y DelaCalle, M. (2018). The effectiveness of a head-heart-hands model for natural and environmental science learning in urban schools. *Evaluation and Program Planning*, 66, 53-62. <https://doi.org/10.1016/j.evalprogplan.2017.09.001>
- Jagannathan, R., Camasso, M. J., y DelaCalle, M. (2019). Promoting cognitive and soft skills acquisition in a disadvantaged public school system: Evidence from the Nurture thru Nature randomized experiment. *Economics of Education Review*, 70, 173-191. <https://doi.org/10.1016/j.econedurev.2019.04.005>
- Jamali, S. M., Ale Ebrahim, N., y Jamali, F. (2022). The role of STEM Education in improving the quality of education: A bibliometric study. *International Journal of Technology and Design Education*, 33(3), 819-840. <https://doi.org/10.1007/s10798-022-09762-1>
- Jiménez, M. (2000). Competencia social: Intervención preventiva en la escuela. *Infancia y Sociedad*, 24, 21-48.
- Kalton, G. (2020). *Introduction to survey sampling*. Sage Publications. <https://doi.org/10.4135/9781071909812.n5>
- Kearney, C. (2016). *Efforts to increase students' interest in pursuing Science, Technology, Engineering and Mathematics studies and careers. National measures taken by 30 countries - 2015 Report*. European Schoolnet.
- Li, Y., Xiao, Y., Wang, K., Zhang, N., Pang, Y., Wang, R., Qi, C., Yuan, Z., Xu, J., Nite, S. B., y Star, J. R. (2022). A systematic review of high impact empirical studies in STEM education. *International Journal of STEM Education*, 9, art. 72. <https://doi.org/10.1186/s40594-022-00389-1>
- National Research Council (2011). *Successful K-12 STEM Education. Identifying effective approaches in Science, Technology, Engineering, and Mathematics*. The National Academies Press.
- National Science Board (2022). *Science and Engineering Indicators*. National Science Foundation.
- Organización para la Cooperación y el Desarrollo Económico (OCDE). (2021). *OECD main science and technology indicatorsOECD main science and technology indicators. RyD Highlights in the March 2021 Publication*. OECD Directorate for Science, Technology and Innovation.
- Organización para la Cooperación y el Desarrollo Económico (OCDE). (2024). *Education at a Glance 2024: OECD Indicators*. OECD Publishing. <https://doi.org/10.1787/c00cad36-en>
- Pintrich, P. R., Smith, D., García, T., y McKeachie, W. (1991). *A manual for the use of the Motivated Strategies for Learning Questionnaire (MSLQ)*. National Center for Research to Improve Postsecondary Teaching and Learning. University of Michigan.
- Pintrich, P. R., Smith, D., García, T., y McKeachie, W. (1993). Reliability and predictive validity of the Motivated Strategies for Learning Questionnaire (MSLQ). *Educational and Psychological Measurement*, 53(3), 801-813. <https://doi.org/10.1177/0013164493053003024>
- Roces, C., Tourón, J., y González, M. C. (1995). Validación preliminar del CEAM II. *Psicológica*, 16, 347-366.
- Roehrig, G. H., Dare, E. A., Ellis, J. A., y Ring-Whalen, E. (2021). Beyond the basics: A detailed conceptual framework of integrated STEM. *Disciplinary and Interdisciplinary Science Education Research*, 3, art. 11. <https://doi.org/10.1186/s43031-021-00041-y>
- Sawilowsky, S. S. (2009). New effect size rules of thumb. *Journal of Modern Applied Statistical Methods*, 8(2), 597-599.
- Shah, A. M., Wylie, C., Gitomer, D., y Noam, G. (2018). Improving STEM program quality in out-of-school-time: Tool development and validation. *Science Education*, 102(2), 38-259. <https://doi.org/10.1002/sce.21327>
- Smith, E., y White, P. (2019). Where Do All the STEM Graduates Go? Higher Education, the Labour Market and Career Trajectories in the UK. *Journal of Science Education and Technology*, 28, 26-40. <https://doi.org/10.1007/s10956-018-9741-5>
- Thibaut, L., Ceuppens, S., De Loof, H., De Meester, J., Goovaerts, L., Struyf, A., Boeve-de Pauw, J., Dehaene, W., Deprez, J., De Cock, M., Hellinckx, L., Knipprath, H., Langie, G., Struyven, K., Van de Velde, D., Van Petegem, P., y Depaepe, F. (2018). Integrated STEM education: A systematic review of instructional practices in secondary education. *European Journal of STEM Education*, 3(1), art. 2. <https://doi.org/10.20897/ejsteme/85525>
- Thurston, A., Cockerill, M., y Chiang, T.-H. (2021). Assessing the differential effects of peer tutoring for tutors and tutees. *Education Sciences*, 11(3), art. 97. <https://doi.org/10.3390/educsci11030097>
- White, P., y Smith, E. (2022). From subject choice to career path: Female STEM graduates in the UK labour market. *Oxford Review of Education*, 28(6), 693-709. <https://doi.org/10.1080/03054985.2021.2011713>
- Whitehead, A. L., Julius, S. A., Cooper, C. L., y Campbell, M. J. (2016). Estimating the sample size for a pilot randomised trial to minimise the overall trial sample size for the external pilot and main trial for a continuous outcome variable. *Statistical Methods in Medical Research*, 25(3), 1057-1073. <https://doi.org/10.1177/0962280215588241>
- Young, J., Ortiz, N., y Young, J. (2017). STEMulating interest: A meta-analysis of the effects of out-of-school time on students STEM interest. *International Journal of Education in Mathematics, Science and Technology*, 5(1), 62-74. <https://doi.org/10.18404/ijemst.61149>
- Zhang, D., y Campbell, T. (2011). The psychometric evaluation of a three-dimension elementary science attitude survey. *Journal of Science Teacher Education*, 22(7), 595-612. <https://doi.org/10.1007/s10972-010-9202-3>
- Zhou, S., Dong, Z., Wang, H. H., y Chiu, M. M. (2024). A meta-analysis of STEM integration on student academic achievement. *Research in Science Education*, 55, 1273-1302. <https://doi.org/10.1007/s11165-024-10216-y>