RICESO

Revista Interdisciplinaria de Ciencias de la Educación, Salud y Sociología

Impacto de un programa de estrategias neurodidácticas en el desarrollo de competencias matemáticas: un estudio preexperimental con medidas pretest y postest en bachillerato

Elian Jamil Hernández Cueva
Evelyn Michelle Ordóñez Merino
Walter Darío Ramón Namcela
Jessica Rosalía Espinosa González

https://doi.org/10.66136/bekryb54

Universidad Nacional de Loja; Ministerio de Educación, Deporte y Cultura del Ecuador.

Ecuador

Impacto de un programa de estrategias neurodidácticas en el desarrollo de competencias matemáticas: un estudio preexperimental con medidas pretest y postest en bachillerato

	Elian Jamil Hernández Cueva **Evelyn Michelle Ordóñez Merino** **Walter Darío Ramón Namcela** **Jessica Rosalía Espinosa González**
	https://orcid.org/0009-0000-1007-9929 https://orcid.org/0009-0006-2873-7436 https://orcid.org/0009-0002-8784-5147 https://orcid.org/0009-0008-7860-8530
	elian.hernandez@unl.edu.ec evelynm.ordonez@docentes.educacion.edu.ec walter.ramon@educacion.gob.ec jessica.espinosa@docentes.educacion.edu.ec
	Universidad Nacional de Loja; Ministerio de Educación, Deporte y Cultura del Ecuador.
	Ecuador
	https://doi.org/10.66136/bekryb54

Resumen

El estudio evaluó el efecto de un programa de estrategias neurodidácticas sobre las competencias matemáticas de estudiantes de tercero de bachillerato. Se adoptó un enfoque cuantitativo con diseño preexperimental de un solo grupo y medidas pretest-postest, en una institución fiscal de Loja, Ecuador. De una población de 840 estudiantes se conformó, mediante muestreo no probabilístico por conveniencia con grupos intactos, una muestra de 128 participantes (69 hombres y 59 mujeres) de cuatro paralelos. El instrumento fue una prueba calificada sobre 20 puntos, estructurada en tres dimensiones: problemas de cantidad, de regularidad y equivalencia, y de forma y movimiento. Durante doce semanas el programa combinó recuperación activa, práctica espaciada e intercalada, regulación emocional del error y representaciones multisensoriales de los objetos matemáticos. El análisis empleó estadística descriptiva, la prueba t para muestras relacionadas y la prueba de rangos con signo de Wilcoxon, incorporada ante la marcada asimetría de la distribución inicial. En la entrada, el 95,3 % se ubicó en el nivel Inicio; en la salida, el 65,6 % alcanzó el nivel Logrado y el Inicio bajó al 1,6 %. La media pasó de 7,33 a 13,49 puntos y la desviación estándar bajó de 1,59 a 1,37, con diferencias estadísticamente significativas, $t(127) = 33,45$; $p < ,001$; $d = 4,16$, confirmadas por el contraste no paramétrico ($Z = 9,81$; $p < ,001$; $r = ,61$). El mayor avance correspondió a regularidad y equivalencia. Los resultados indican mejoras sustantivas, aunque la ausencia de grupo de control impide una atribución causal concluyente.

Palabras claves: neuroeducación; competencias matemáticas; educación secundaria; funciones ejecutivas; práctica espaciada.

Received: 04/09/2026
Accepted: 23/09/2026
Published: 25/09/2026

Revista Interdisciplinaria de Ciencias de la Educación, Salud y Sociología
https://www.riceso.org

editor@riceso.org

© 2026. Este artículo es un documento de acceso abierto distribuido bajo los términos y condiciones de la **Licencia Creative Commons, Atribución-NoComercial 4.0 Internacional.**

Impact of a neurodidactic strategies programme on the development of mathematical competencies: a pre-experimental pretest-posttest study in high school

Abstract

The study evaluated the effect of a program of neurodidactic strategies on the mathematical competencies of third-year upper secondary students. A quantitative approach was adopted, using a one-group pre-experimental design with pretest-posttest measures, in a public institution in Loja, Ecuador. From a population of 840 students, a sample of 128 participants (69 male, 59 female) from four intact class groups was selected through non-probabilistic convenience sampling. The instrument was a mathematical competency test scored out of 20 points and organized into three dimensions: problem solving involving quantity, regularity and equivalence, and shape and movement. Over twelve weeks, the program combined active recall, spaced and interleaved practice, emotional regulation of error, and multisensory representations of mathematical objects. Data were analyzed using descriptive statistics, the paired-samples Student t test, and the Wilcoxon signed-rank test, added in response to the marked skewness of the initial distribution. At pretest, 95.3% of students were placed at the Initial level; at posttest, 65.6% reached the Achieved level and the Initial level fell to 1.6%. The mean rose from 7.33 to 13.49 points and the standard deviation decreased from 1.59 to 1.37, with statistically significant differences, $t(127) = 33.45$, $p < .001$, $d = 4.16$, confirmed by the nonparametric contrast ($Z = 9.81$, $p < .001$, $r = .61$). The largest gain corresponded to the regularity and equivalence dimension. The findings indicate substantive improvement, although the absence of a control group precludes a conclusive causal attribution.

Keywords: neuroeducation; mathematical competencies; secondary education; executive functions; spaced practice.

Introducción

La enseñanza de la matemática en el bachillerato arrastra un problema que ninguna reforma curricular reciente ha logrado resolver: los estudiantes egresan sabiendo ejecutar procedimientos, pero no sabiendo decidir cuándo aplicarlos. Esa distancia entre la ejecución algorítmica y el razonamiento con sentido es la que motiva el presente trabajo. La neurodidáctica, entendida como el conjunto de decisiones instruccionales que incorporan evidencia de las ciencias cognitivas y de la neurociencia del aprendizaje, ofrece un marco para intervenir sobre esa distancia. Conviene, sin embargo, ser precisos sobre su alcance: la neurociencia no prescribe métodos de aula, sino que restringe el espacio de hipótesis plausibles sobre cómo aprende un estudiante y ayuda a descartar prácticas que carecen de fundamento (Leisman, 2023). Esa cautela es la que orienta el diseño que aquí se reporta.

El desempeño matemático de la región latinoamericana ofrece el telón de fondo del problema. Los resultados de la evaluación internacional más reciente muestran una caída sin precedentes en el promedio de los países de la OCDE, con un descenso de quince puntos en matemática entre 2018 y 2022, y sitúan a la región muy por debajo de ese promedio, con tres de cada cuatro estudiantes por debajo del nivel mínimo de competencia (OCDE, 2023). Ecuador no participa en ese estudio, de modo que la referencia nacional más pertinente proviene del Estudio Regional Comparativo y Explicativo, en el que el país obtuvo puntajes superiores al promedio regional en matemática de cuarto y séptimo de educación general básica, aunque con una proporción todavía considerable de estudiantes en el nivel de desempeño más bajo (UNESCO-OREALC & LLECE, 2021). Ese contraste es relevante: el problema ecuatoriano no se explica solo por rezago respecto de la región, sino por la dificultad de convertir logros tempranos en competencia sostenida al final de la trayectoria escolar.

Desde la investigación cognitiva, el sustrato de esa dificultad está razonablemente documentado. Un metaanálisis de estudios longitudinales con más de cien mil participantes estableció una asociación moderada y estable entre funciones ejecutivas y habilidades matemáticas a lo largo del desarrollo, con la memoria de trabajo como componente de mayor peso (Tette et al., 2026). En muestras de primaria y secundaria, distintas combinaciones de inhibición, flexibilidad cognitiva y memoria de trabajo predicen tanto el rendimiento como la

dificultad matemática, y lo hacen de manera específica para cada grado escolar (Iglesias-Sarmiento et al., 2023). La metacognición aporta una vía complementaria: en adolescentes de once a dieciséis años, la capacidad de monitorear y regular el propio proceso de resolución se asocia positivamente con el desempeño matemático (Muncer et al., 2022). Estos hallazgos sugieren que una intervención dirigida al bachillerato debería actuar sobre el control ejecutivo y la autorregulación, y no únicamente sobre el contenido curricular.

El nivel neural añade matices que importan para el diseño instruccional. El surco intraparietal sostiene la representación de la magnitud numérica, y estudios de resonancia magnética funcional muestran que sus patrones de activación distinguen tareas y formatos numéricos, es decir, que el cerebro no trata del mismo modo un numeral arábigo, una configuración de dedos o un enunciado verbal (Koch et al., 2023). A ello se suma evidencia de que rasgos morfológicos tempranos de esa misma región se relacionan con las habilidades numéricas simbólicas, lo que introduce una restricción a la idea de una plasticidad ilimitada (Roell et al., 2021). La lectura pedagógica razonable no es que el desempeño esté determinado, sino que la variación entre estudiantes tiene una base biológica parcial y que la instrucción debe ofrecer múltiples formatos de acceso al mismo concepto.

La dimensión afectiva es el segundo eje del problema. Un metaanálisis de amplio alcance confirmó una correlación negativa pequeña a moderada entre ansiedad matemática y rendimiento, sostenida a través de muestras y de instrumentos de medida (Barroso et al., 2021). Lo relevante es que esa relación es modificable: las intervenciones dirigidas a la ansiedad matemática producen efectos moderados tanto en la reducción del malestar como en la mejora del desempeño, y son más eficaces cuando combinan apoyo cognitivo y regulación emocional (Sammallahti et al., 2023). Una síntesis posterior, centrada en programas escolares con estudiantes de educación básica y media, llegó a conclusiones convergentes respecto de la conveniencia de integrar componentes terapéuticos y de destreza (Codding et al., 2023). Estos resultados justifican que un programa neurodidáctico incorpore el tratamiento del error y la regulación de la respuesta emocional como componentes instruccionales, y no como añadidos motivacionales.

Un tercer eje corresponde a las estrategias de estudio con mayor respaldo empírico. La práctica espaciada y la recuperación activa constituyen, junto con la práctica intercalada, el

conjunto de técnicas cuya eficacia ha sido replicada con mayor consistencia (Carpenter et al., 2022). La evidencia experimental muestra que espaciar los episodios de recuperación mejora la retención frente a la práctica masiva (Latimier et al., 2021) y que las pruebas breves aplicadas en el aula elevan el aprendizaje de manera robusta (Yang et al., 2021). En el dominio matemático específicamente, una síntesis reciente encontró un efecto pequeño a medio de la práctica espaciada frente a la masiva, con magnitudes menores a las observadas en otros dominios y sin evidencia concluyente de un efecto sistemático de la recuperación cuando se la contrasta con el reestudio (Murray et al., 2025). Estudios de aula en matemática habían mostrado antes que espaciar la práctica mejora los puntajes y reduce el exceso de confianza de los estudiantes (Emeny et al., 2021), y que el beneficio del espaciamiento se extiende a otros dominios de aprendizaje escolar (Kim & Webb, 2022). Esta literatura obliga a moderar las expectativas: el aporte de estas técnicas en matemática es real pero acotado, y su rendimiento depende del modo en que se articulen con la carga cognitiva de la tarea.

La teoría de la carga cognitiva ofrece precisamente ese marco de articulación. Su formulación actual sostiene que la instrucción debe ajustarse a los límites de la memoria de trabajo cuando se procesa información biológicamente secundaria, categoría a la que pertenece la matemática escolar (Sweller, 2022). Revisiones posteriores del propio programa teórico reconocen la necesidad de incorporar aportes de otras tradiciones y de someter sus predicciones a estándares de replicación más exigentes (Sweller, 2023), y el debate abierto sobre la combinación de instrucción directa e indagación muestra que la cuestión dista de estar cerrada (Sweller et al., 2024). Para el diseño de una intervención en bachillerato, la implicación práctica es doble: conviene apoyar la resolución de problemas con estructuras de guía explícita y, al mismo tiempo, retirar progresivamente esa guía a medida que se consolidan esquemas.

A este cuadro se suma un fenómeno reciente que altera las condiciones del aula. La disponibilidad de asistentes conversacionales de inteligencia artificial ha modificado el modo en que los estudiantes enfrentan una consigna matemática, al permitir trasladar el esfuerzo de resolución a un sistema externo. Un estudio con más de seiscientos participantes documentó una asociación negativa entre el uso frecuente de estas herramientas y las puntuaciones en pensamiento crítico, con la descarga cognitiva actuando como mediador, y encontró una dependencia mayor en los grupos de menor edad (Gerlich, 2025). Desde la perspectiva de la

cognición distribuida, el problema no es la delegación en sí, sino la delegación indiscriminada de los procesos que constituyen el objetivo del aprendizaje (Grinschgl & Neubauer, 2022). Esta distinción es la que convierte a la recuperación activa en una estrategia especialmente pertinente en el contexto actual: obliga al estudiante a producir la respuesta antes de disponer de ella.

Existe, por último, una advertencia que la literatura sobre neuroeducación ha vuelto ineludible. Los neuromitos, entendidos como interpretaciones erróneas de hallazgos neurocientíficos aplicadas a la enseñanza, siguen presentando una prevalencia elevada entre docentes en ejercicio y en formación (Novak-Geiger, 2023). Un ensayo con 807 docentes de primaria y secundaria mostró que una formación en línea de quince horas reduce de manera considerable y duradera esas creencias, lo que indica que el problema es corregible mediante intervenciones de bajo costo (Ruiz-Martín et al., 2022). La consecuencia metodológica para este estudio es directa: un programa que se declare neurodidáctico debe explicitar qué evidencia sostiene cada uno de sus componentes y evitar cualquier apelación a estilos de aprendizaje, dominancia hemisférica o fórmulas de estimulación cerebral sin respaldo.

La revisión precedente permite identificar el vacío que este trabajo aborda. La evidencia sobre funciones ejecutivas, ansiedad matemática y estrategias de práctica proviene mayoritariamente de contextos anglosajones y europeos, y se concentra en educación primaria o universitaria. Son escasos los estudios que documentan, con datos de aula y en instituciones públicas de la región andina, el comportamiento de un programa que integre estos componentes en el tramo final del bachillerato, justo donde se juega la transición hacia la educación superior. El presente estudio ofrece esa documentación.

En consecuencia, el objetivo general fue evaluar el efecto de un programa de estrategias neurodidácticas sobre el desarrollo de competencias matemáticas en estudiantes de tercero de bachillerato de una institución fiscal de Loja. Como objetivos específicos se plantearon: caracterizar el nivel de logro del estudiantado antes y después de la intervención; comparar el desempeño en las tres dimensiones evaluadas; y estimar la magnitud del cambio observado. Se formularon dos hipótesis. La hipótesis nula (H0) establece que no existen diferencias estadísticamente significativas entre las puntuaciones de entrada y de salida en competencias matemáticas. La hipótesis alterna (H1) sostiene que las puntuaciones de salida son significativamente superiores a las de entrada tras la aplicación del programa.

Metodología

Enfoque y diseño

La investigación se inscribe en un paradigma postpositivista y adopta un enfoque cuantitativo. El diseño corresponde a un estudio preexperimental de un solo grupo con medidas pretest y postest, de corte longitudinal corto. Se emplea aquí la denominación preexperimental, y no cuasi-experimental, por una razón técnica que conviene explicitar: un diseño cuasi-experimental requiere al menos un grupo de comparación no equivalente, condición que este estudio no cumple. Esta precisión, ausente en versiones preliminares del trabajo, es determinante para interpretar correctamente los resultados y se retoma en el apartado de limitaciones. La variable independiente es la aplicación del programa de estrategias neurodidácticas; la variable dependiente es el nivel de competencias matemáticas medido por la prueba descrita más adelante.

Participantes y contexto

El estudio se desarrolló en una unidad educativa de sostenimiento fiscal ubicada en la parroquia El Sagrario, cantón Loja, provincia de Loja (Zona 7 de planificación, régimen Sierra), que atiende jornadas matutina, vespertina y nocturna. La institución cuenta con una planta docente de 54 profesionales y una población estudiantil de 840 estudiantes, de la cual se extrajo la muestra.

La selección se realizó mediante muestreo no probabilístico por conveniencia, trabajando con grupos intactos, e incluyó a la totalidad de estudiantes matriculados en tercero de bachillerato distribuidos en los paralelos A, B, C y D (n = 128). El grupo estuvo conformado por 69 hombres (53,9 %) y 59 mujeres (46,1 %), con una edad media de 17,2 años. Cada paralelo reunió a 32 estudiantes, con una composición por sexo desigual entre aulas: el paralelo A registró 18 hombres y 14 mujeres; el B, 12 hombres y 20 mujeres; el C, 15 hombres y 17 mujeres; y el D, 24 hombres y 8 mujeres. Esta heterogeneidad se documenta porque, al tratarse de grupos intactos, no fue posible equilibrarla mediante asignación. Debe subrayarse que una muestra obtenida por este procedimiento no es representativa de la población en sentido estadístico, por lo que los resultados no admiten generalización más allá del grupo estudiado. Se consideraron criterios de inclusión la matrícula vigente en tercero de bachillerato y la participación en ambas mediciones;

se excluyeron los casos con inasistencia a alguna de las dos aplicaciones. No se registraron pérdidas entre el pretest y el postest, de modo que los 128 casos aportan medidas completas.

Instrumento

Se utilizó una prueba de competencias matemáticas de elaboración propia, construida a partir de los estándares curriculares vigentes para el nivel y puntuada sobre 20 puntos. El instrumento se organiza en tres dimensiones: resolución de problemas de cantidad (D1), resolución de problemas de regularidad y equivalencia (D2) y resolución de problemas de forma y movimiento (D3). Las puntuaciones se interpretan mediante cuatro niveles de logro: Inicio, En proceso, Logrado y Destacado, definidos por puntos de corte establecidos a priori sobre la escala total y sobre cada dimensión.

La validez de contenido y de constructo se estableció mediante juicio de expertos ciegos, con la participación de 5 especialistas en didáctica de la matemática y evaluación educativa. La consistencia interna se estimó sobre la aplicación de entrada mediante el coeficiente alfa de Cronbach, que alcanzó un valor de ,87 para la escala total, indicativo de una fiabilidad alta. Se aplicaron formas paralelas en el pretest y el postest para reducir el efecto de familiarización con los ítems. Durante la fase de intervención se emplearon, además, guías de observación sistemática de aula para registrar la fidelidad de implementación de cada componente.

Programa de intervención

El programa se organizó en torno a cuatro componentes, cada uno anclado a un cuerpo de evidencia identificable. El primero consistió en recuperación activa: cada sesión se abrió con una serie breve de preguntas sobre contenidos trabajados en sesiones anteriores, resueltas sin acceso a apuntes ni a dispositivos, con retroalimentación inmediata (Carpenter et al., 2022; Yang et al., 2021). El segundo componente fue la práctica espaciada e intercalada: los ejercicios de cada bloque temático se distribuyeron a lo largo de varias semanas y se mezclaron con problemas de bloques previos, de modo que dos ejercicios consecutivos no requirieran la misma estrategia de resolución (Emeny et al., 2021; Murray et al., 2025). El tercero abordó la regulación emocional del error, mediante rutinas de análisis público de respuestas incorrectas orientadas a identificar el punto exacto de la falla en el razonamiento, con el propósito de reducir la respuesta de evitación asociada a la ansiedad matemática (Codding et al., 2023; Sammallahti et al., 2023). El cuarto

consistió en la representación multisensorial de los objetos matemáticos, presentando cada concepto en formato simbólico, gráfico, verbal y manipulativo dentro de la misma secuencia didáctica, en coherencia con la evidencia sobre el procesamiento diferenciado de formatos numéricos (Koch et al., 2023).

La secuencia instruccional siguió un principio de retirada progresiva de la guía: los problemas nuevos se introdujeron con ejemplos resueltos y, a medida que se consolidaba el esquema de resolución, se sustituyeron por problemas de resolución autónoma, de acuerdo con las prescripciones de la teoría de la carga cognitiva para material biológicamente secundario (Sweller, 2022, 2023). Se descartaron de manera deliberada los componentes carentes de respaldo empírico que suelen asociarse a la etiqueta neurodidáctica, en particular la adaptación de la enseñanza a supuestos estilos de aprendizaje y las actividades de lateralización hemisférica (Ruiz-Martín et al., 2022). El programa se desarrolló a lo largo de doce semanas consecutivas, con tres sesiones semanales de 45 minutos cada una, lo que equivale a 36 sesiones y 27 horas de trabajo efectivo dentro de la carga horaria ordinaria de la asignatura. La implementación estuvo a cargo de tres docentes investigadores, quienes trabajaron con una planificación común y sesiones de coordinación semanal para asegurar la equivalencia del tratamiento entre paralelos.

Estructura de la sesión y secuencia semanal

Las 36 sesiones compartieron una arquitectura interna fija. Esa invariancia fue deliberada: cuando la rutina de trabajo es predecible, el estudiante no gasta recursos de memoria de trabajo en averiguar qué se espera de él y puede destinarlos al contenido matemático, que es donde la carga resulta pertinente (Sweller, 2022). La Tabla 1 detalla la distribución del tiempo dentro de la sesión tipo.

Tabla 1

Estructura de la sesión tipo del programa de intervención (45 minutos)

Min.	Momento	Descripción de la actividad	Componente
0-8	Apertura acumulativa	Serie de tres a cinco preguntas breves sobre contenidos de sesiones anteriores, resueltas por escrito y de forma individual, sin apuntes ni dispositivos, con verificación inmediata en el pizarrón.	Recuperación activa
8-15	Análisis del error	Revisión colectiva de una respuesta incorrecta frecuente, orientada a localizar el punto exacto del razonamiento en que se produce la falla, y no a señalar al autor de la respuesta.	Regulación emocional
15-27	Presentación multiformato	Introducción del objeto matemático de la sesión en cuatro registros sucesivos: simbólico, gráfico, verbal y manipulativo o dinámico.	Representación multisensorial
27-40	Práctica con guía decreciente	Secuencia de tres problemas: ejemplo íntegramente resuelto, problema de completar con pasos parcialmente omitidos y problema de resolución autónoma.	Carga cognitiva
40-45	Cierre intercalado	Dos problemas de bloques temáticos anteriores, presentados de modo que ningún par consecutivo comparta estrategia de resolución.	Práctica espaciada e intercalada

Nota. Elaboración propia. La suma de los tiempos corresponde a un periodo de clase de 45 minutos. Los rangos indican el momento de inicio y fin de cada bloque dentro de la sesión.

La distribución semanal siguió una lógica de tres sesiones con funciones diferenciadas. La primera sesión de cada semana introdujo el contenido nuevo con la máxima densidad de guía; la segunda concentró la práctica intercalada y el trabajo sobre el error; la tercera se destinó a recuperación acumulativa y a problemas que exigían transferir el procedimiento a una situación no vista. Esta asignación permitió que el intervalo entre la presentación de un contenido y su

primera recuperación espaciada fuese de al menos 48 horas, condición que la evidencia sobre el efecto de espaciamiento identifica como relevante para la retención (Latimier et al., 2021).

La secuencia de las doce semanas se organizó en cuatro fases. Las dos primeras semanas cumplieron una función de encuadre y de alfabetización estratégica: antes de aplicar las técnicas, se explicó al estudiantado por qué recuperar información de memoria produce más aprendizaje que releerla, dado que estas estrategias resultan contraintuitivas y su adopción depende de que el aprendiz comprenda su fundamento (Carpenter et al., 2022). Las tres fases siguientes se dedicaron sucesivamente a las dimensiones D1, D2 y D3, con una proporción creciente de práctica acumulada procedente de las dimensiones ya trabajadas. La semana final se reservó a problemas de integración y a la aplicación del postest. La Tabla 2 presenta la secuencia completa.

Tabla 2

Secuencia semanal del programa de intervención

Sem.	Fase	Contenido matemático	Énfasis estratégico	Práctica corriente / acumulada
1	Encuadre	Aplicación del pretest y presentación del acuerdo de trabajo del aula	Establecimiento de la rutina de sesión	-
2	Alfabetización estratégica	Repaso de operatoria elemental y lenguaje algebraico	Fundamento de la recuperación activa y del tratamiento del error	100 / 0
3	D1. Cantidad	Números reales: orden, aproximación y notación científica	Presentación multiformato de la magnitud numérica	100 / 0
4	D1. Cantidad	Razones, proporcionalidad, porcentajes e interés	Ejemplos resueltos y recuperación acumulativa	70 / 30
5	D1. Cantidad	Problemas contextualizados de estimación y cálculo	Problemas de completar; retirada parcial de la guía	60 / 40

Sem.	Fase	Contenido matemático	Énfasis estratégico	Práctica corriente / acumulada
6	D2. Regularidad	Patrones numéricos y sucesiones	Discriminación de la estructura antes de elegir estrategia	60 / 40
7	D2. Regularidad	Ecuaciones e inecuaciones de primer grado	Práctica intercalada intensiva	50 / 50
8	D2. Regularidad	Sistemas de ecuaciones y noción de función	Traducción entre registros de representación	50 / 50
9	D3. Forma y movimiento	Figuras planas: perímetro, área y semejanza	Trabajo con material manipulativo y representación dinámica	50 / 50
10	D3. Forma y movimiento	Cuerpos geométricos, superficie y volumen	Visualización y rotación mental de objetos	40 / 60
11	D3. Forma y movimiento	Transformaciones en el plano y localización	Problemas no rutinarios de resolución autónoma	40 / 60
12	Integración y salida	Problemas que combinan las tres dimensiones; aplicación del postest	Autonomía plena, sin guía instruccional	0 / 100

Nota. Elaboración propia. D1 = resolución de problemas de cantidad; D2 = resolución de problemas de regularidad y equivalencia; D3 = resolución de problemas de forma y movimiento. La última columna expresa el porcentaje aproximado de ejercicios correspondientes al contenido de la semana en curso frente al de semanas anteriores.

Los materiales consistieron en cuadernillos impresos de problemas, un banco de ejemplos resueltos por dimensión, material manipulativo para el trabajo geométrico y un entorno de geometría dinámica de acceso libre empleado en las sesiones de la fase D3. No se utilizaron plataformas comerciales de entrenamiento cerebral ni programas de estimulación cognitiva computarizada, cuya eficacia para la transferencia al rendimiento académico continúa siendo objeto de debate (Birtwistle et al., 2025).

La fidelidad de implementación se controló mediante guías de observación sistemática aplicadas por un observador externo en 12 de las 36 sesiones, seleccionadas de forma aleatoria, lo que representa un tercio del total impartido. La guía registró la presencia y la duración efectiva de cada uno de los cinco bloques de la sesión tipo, así como la proporción real de práctica acumulada frente a la planificada. El índice de fidelidad global alcanzó un 98 %, calculado como el cociente entre los componentes ejecutados conforme a lo previsto y el total de componentes observados, valor que respalda la correspondencia entre el programa planificado y el efectivamente aplicado.

Procedimiento

El pretest se aplicó durante la primera semana del periodo lectivo, en horario de clase y bajo condiciones estandarizadas de administración. A continuación se desarrolló la intervención. El postest se aplicó en la semana siguiente a la finalización del programa, con la forma paralela del instrumento y en condiciones equivalentes de aplicación. La calificación fue realizada por evaluadores que desconocían el momento de medición al que correspondía cada protocolo, con el fin de reducir el sesgo de expectativa.

Análisis de datos

Los datos se procesaron con el paquete estadístico IBM SPSS Statistics, versión 29. El análisis descriptivo comprendió frecuencias absolutas y relativas por nivel de logro, medidas de tendencia central (media y mediana), de dispersión (desviación estándar, varianza y coeficiente de variación) y de forma (asimetría). El supuesto de normalidad se evaluó mediante la prueba de Kolmogorov-Smirnov y mediante la razón entre el coeficiente de asimetría y su error típico. Ambos procedimientos arrojaron resultados divergentes: la prueba de Kolmogorov-Smirnov no permitió rechazar la normalidad ($p > ,05$), mientras que la razón crítica de asimetría del pretest excedió con amplitud el umbral convencional. Ante esa divergencia se optó por reportar los dos contrastes posibles. El principal fue la prueba t de Student para muestras relacionadas, aplicable porque el supuesto que esta prueba requiere se refiere a la distribución de las diferencias y no a la de las puntuaciones brutas. Como verificación de robustez se aplicó la prueba no paramétrica de rangos con signo de Wilcoxon. En ambos casos se trabajó con un nivel de significación de $\alpha = ,05$.

La magnitud del efecto se estimó mediante la d de Cohen para medidas repetidas, obtenida a partir de la desviación estándar combinada de ambos momentos, y se complementó con el coeficiente r de Rosenthal, calculado como el cociente entre el estadístico Z y la raíz cuadrada del número total de observaciones. Se reportan intervalos de confianza al 95 % para la diferencia de medias. La presentación de resultados estadísticos, tablas y figuras sigue las prescripciones de la séptima edición del manual de publicaciones de la American Psychological Association (American Psychological Association, 2020). Conviene precisar que la d de Cohen se calculó sobre la desviación estándar combinada de ambos momentos y no sobre la desviación estándar de las diferencias. Calculada de este segundo modo, la magnitud del efecto desciende a 2,96, valor que sigue siendo muy elevado pero considerablemente inferior al primero. Se reportan ambas estimaciones porque la elección del denominador altera la lectura del resultado y esa decisión debe quedar explícita.

Consideraciones éticas

El estudio contó con la autorización de las autoridades institucionales y con el aval del comité de ética de la institución. Se obtuvo el consentimiento informado de los representantes legales y el asentimiento informado de los estudiantes, dado que una parte de la muestra corresponde a personas menores de edad. Los datos fueron anonimizados mediante códigos numéricos antes del procesamiento, y la participación fue voluntaria, con posibilidad de retiro en cualquier momento sin consecuencias académicas. Al finalizar el estudio se socializaron los resultados agregados con la comunidad educativa.

Resultados

El análisis del desempeño global antes y después de la intervención muestra un desplazamiento sustantivo en la distribución de los niveles de logro. La Tabla 3 presenta las frecuencias y los porcentajes correspondientes a ambos momentos de medición.

Tabla 3

Distribución del rendimiento global en competencias matemáticas en el pretest y el postest (n = 128)

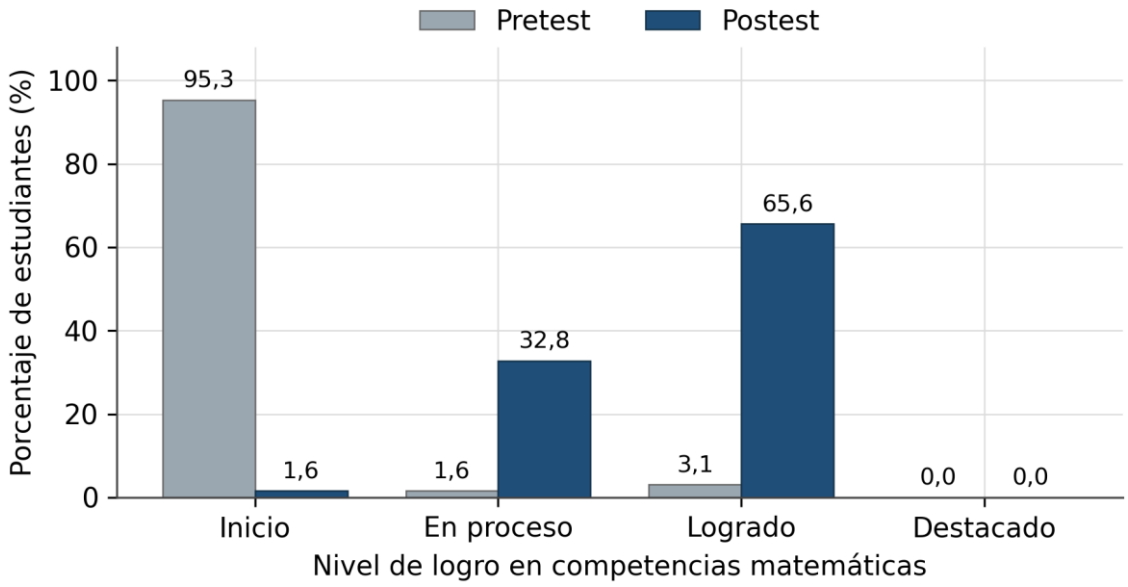
Nivel de logro	Pretest f	Pretest %	Postest f	Postest %
Inicio	122	95,3	2	1,6
En proceso	2	1,6	42	32,8
Logrado	4	3,1	84	65,6
Destacado	0	0,0	0	0,0
Total	**128**	**100,0**	**128**	**100,0**

Nota. Elaboración propia. f = frecuencia absoluta. Los niveles de logro se derivan de los puntos de corte aplicados sobre la puntuación total de la prueba (escala 0–20).

El diagnóstico de entrada revela una situación crítica: 122 de los 128 estudiantes, equivalentes al 95,3 % de la muestra, se ubicaron en el nivel Inicio, y solo cuatro alcanzaron el nivel Logrado. Tras la intervención, 84 estudiantes (65,6 %) se situaron en el nivel Logrado y 42 (32,8 %) en el nivel En proceso, mientras que el nivel Inicio quedó reducido a dos casos (1,6 %). Ningún estudiante alcanzó el nivel Destacado en la puntuación global en ninguno de los dos momentos, resultado que se explica por la exigencia del punto de corte correspondiente, que requiere un desempeño alto y simultáneo en las tres dimensiones. La Figura 1 representa gráficamente ese desplazamiento.

Figura 1

Porcentaje de estudiantes por nivel de logro en competencias matemáticas en el pretest y el postest



Nota. Elaboración propia a partir de los datos recolectados (n = 128).

Al desagregar el desempeño por dimensiones se observa que el avance fue transversal, aunque con magnitudes desiguales. La Tabla 4 detalla la distribución por nivel de logro en cada una de las tres dimensiones evaluadas.

Tabla 4

Distribución del desempeño por dimensión en el pretest y el postest (n = 128)

Dimensión	Nivel	Pretest f	Pretest %	Postest f	Postest %
D1. Cantidad	Inicio	118	92,2	0	0,0
	En proceso	6	4,7	61	47,7
	Logrado	4	3,1	58	45,3
	Destacado	0	0,0	9	7,0
D2. Regularidad y equivalencia	Inicio	118	92,2	3	2,3

Dimensión	Nivel	Pretest f	Pretest %	Postest f	Postest %
	En proceso	9	7,0	33	25,8
	Logrado	1	0,8	91	71,1
	Destacado	0	0,0	1	0,8
D3. Forma y movimiento	Inicio	122	95,3	8	6,2
	En proceso	3	2,3	65	50,8
	Logrado	3	2,3	52	40,6
	Destacado	0	0,0	3	2,3

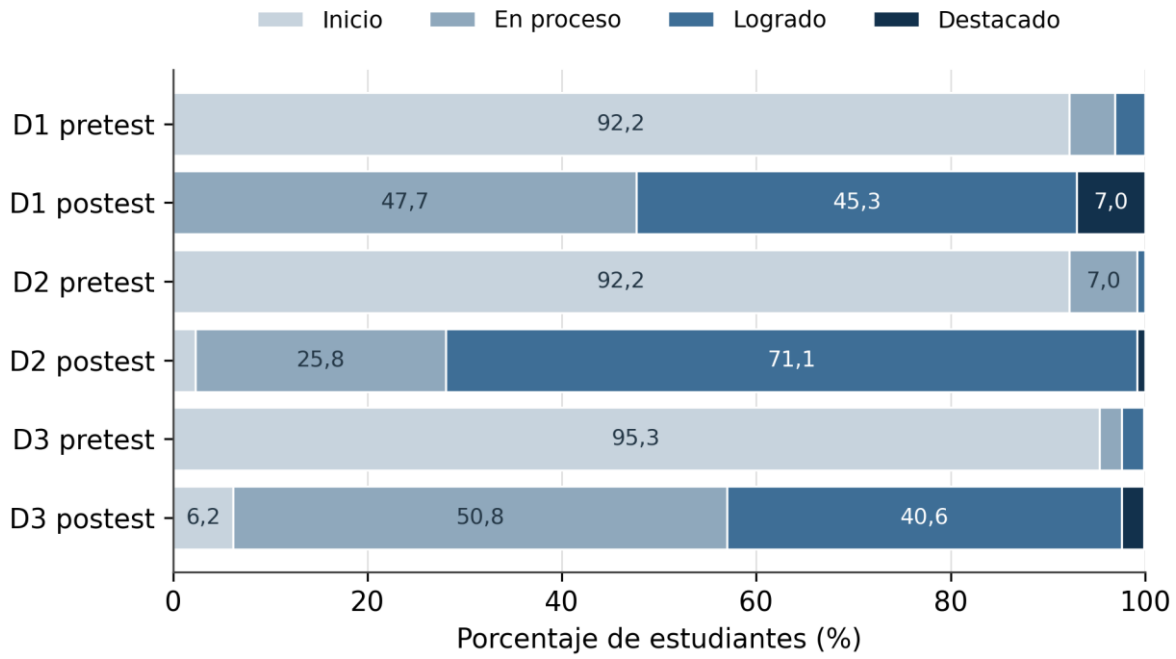
Nota. Elaboración propia. f = frecuencia absoluta. Los porcentajes se calculan sobre el total de la muestra en cada dimensión y momento de medición.

La dimensión de regularidad y equivalencia concentró el mayor porcentaje de estudiantes en el nivel Logrado tras la intervención (71,1 %), seguida por la dimensión de cantidad (45,3 %) y por la de forma y movimiento (40,6 %). El patrón resulta coherente con la naturaleza de las tareas: los problemas de regularidad y equivalencia descansan en el reconocimiento y la extensión de patrones, un tipo de razonamiento que se beneficia de manera directa de la práctica intercalada, ya que esta obliga a discriminar entre estructuras superficialmente similares antes de seleccionar una estrategia. La dimensión de forma y movimiento, en cambio, retuvo la mayor proporción de estudiantes en niveles bajos, con un 6,2 % todavía en Inicio y un 50,8 % en En proceso, lo que sugiere que el razonamiento espacial requiere un tiempo de consolidación mayor que el trabajado en el programa.

El nivel Destacado se mantuvo marginal en las tres dimensiones, con valores comprendidos entre el 0,8 % y el 7,0 %, correspondiendo el valor más alto a la dimensión de cantidad. La lectura razonable de este dato es que el programa desplazó al grueso del estudiantado hacia el umbral de competencia esperado, pero no produjo desempeños de excelencia; alcanzar ese nivel probablemente exige una exposición más prolongada y un trabajo específico sobre problemas de transferencia. La Figura 2 permite comparar visualmente el perfil de las tres dimensiones.

Figura 2

Distribución porcentual del desempeño por dimensión y momento de medición



Nota. Elaboración propia. D1 = resolución de problemas de cantidad; D2 = resolución de problemas de regularidad y equivalencia; D3 = resolución de problemas de forma y movimiento. Las etiquetas se omiten en los segmentos inferiores al 5,5 % para preservar la legibilidad.

La Tabla 5 reúne los estadísticos descriptivos de las puntuaciones totales en ambos momentos. Se han corregido en esta versión los dos valores de varianza, que no correspondían al cuadrado de las desviaciones estándar reportadas, y se han incorporado el coeficiente de variación y el error típico de asimetría, necesarios para interpretar la dispersión y la forma de las distribuciones.

Tabla 5

Estadísticos descriptivos de las puntuaciones totales en el pretest y el postest

Estadístico	Pretest	Postest
n	128	128
Media	7,33	13,49

Estadístico	Pretest	Postest
Desviación estándar	1,59	1,37
Varianza	2,53	1,88
Coeficiente de variación (%)	21,7	10,2
Asimetría	2,51	-0,23
Error típico de asimetría	0,21	0,21
Razón crítica de asimetría	11,73	-1,07

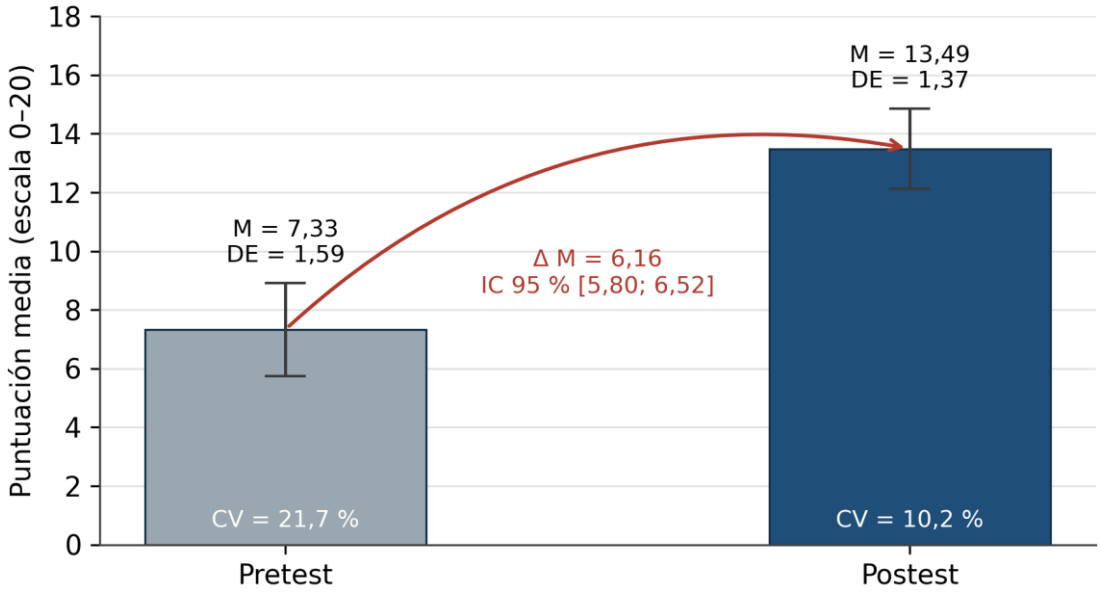
Nota. Elaboración propia. Puntuaciones expresadas sobre una escala de 0 a 20 puntos. El coeficiente de variación se calcula como el cociente entre la desviación estándar y la media, expresado en porcentaje. La razón crítica es el cociente entre el coeficiente de asimetría y su error típico; valores fuera del intervalo ±1,96 indican apartamiento significativo de la simetría.

La media aritmética se incrementó en 6,16 puntos, al pasar de 7,33 a 13,49 sobre 20, con un intervalo de confianza al 95 % para la diferencia de medias comprendido entre 5,80 y 6,52 puntos. La dispersión disminuyó de manera apreciable: la desviación estándar bajó de 1,59 a 1,37 y, sobre todo, el coeficiente de variación se redujo a menos de la mitad, de 21,7 % a 10,2 %. Este segundo indicador resulta más informativo que la desviación estándar por sí sola, dado que las medias de ambos momentos difieren de forma sustancial; su descenso indica que el grupo no solo mejoró, sino que se volvió más homogéneo en términos relativos.

El cambio en la forma de la distribución confirma esa lectura. En el pretest, el coeficiente de asimetría de 2,51, con un error típico de 0,21, produce una razón crítica de 11,73, muy por encima del umbral convencional de ±1,96, lo que indica una distribución fuertemente sesgada hacia la derecha: la gran mayoría de los estudiantes se acumulaba en puntuaciones bajas y unos pocos casos aislados alcanzaban valores altos. En el postest, la asimetría de -0,23 arroja una razón crítica de -1,07, compatible con una distribución aproximadamente simétrica con una leve concentración en los valores superiores. La Figura 3 sintetiza esta comparación.

Figura 3

Comparación de las puntuaciones medias y de la dispersión en el pretest y el postest



Nota. Elaboración propia a partir de los estadísticos descriptivos de la Tabla 5 (n = 128). Las barras de error representan ± 1 desviación estándar. CV = coeficiente de variación. Δ M = diferencia de medias entre ambos momentos.

Contraste de hipótesis

La prueba t de Student para muestras relacionadas arrojó diferencias estadísticamente significativas entre las puntuaciones de entrada (M = 7,33; DE = 1,59) y las de salida (M = 13,49; DE = 1,37), t(127) = 33,45; p < ,001. La verificación no paramétrica mediante la prueba de rangos con signo de Wilcoxon condujo al mismo resultado, Z = 9,81; p < ,001, con un coeficiente r de Rosenthal de ,61, interpretable como un efecto grande. La convergencia de ambos procedimientos permite rechazar la hipótesis nula y aceptar la hipótesis alterna: las puntuaciones de salida son significativamente superiores a las de entrada.

Un dato que conviene explicitar es la desviación estándar de las diferencias implícita en el estadístico t, que asciende a 2,08 puntos. Ese valor supera a las desviaciones estándar de ambas mediciones por separado, lo que implica una correlación prácticamente nula entre el desempeño inicial y el final de cada estudiante. En términos sustantivos significa que la posición relativa del alumnado se reordenó por completo: quienes puntuaban más alto en el pretest no fueron

necesariamente quienes más puntuaron en el postest. Este comportamiento es coherente con el efecto de suelo observado en la medición inicial, donde el 95,3 % de la muestra se concentraba en el nivel más bajo y el instrumento apenas discriminaba entre estudiantes.

El tamaño del efecto calculado sobre la métrica de las medias resulta considerablemente mayor. La d de Cohen para medidas repetidas, obtenida a partir de la desviación estándar combinada de ambos momentos (DE combinada = 1,48), asciende a 4,16. Calculada sobre la desviación estándar de las diferencias, en cambio, desciende a 2,96. Este valor excede en un orden de magnitud las magnitudes habitualmente reportadas en intervenciones educativas, que rara vez superan g = 0,60 incluso en los programas más eficaces (Sammallahti et al., 2023; Murray et al., 2025). Antes que celebrarse como evidencia de una eficacia excepcional, un valor de esta naturaleza debe tratarse como una señal de alerta metodológica: puede reflejar un efecto real de gran tamaño, pero también la confluencia de un diseño sin grupo de control, un efecto de práctica derivado de la doble aplicación del instrumento, una regresión a la media a partir de un piso extremadamente bajo, o una sensibilidad limitada del instrumento en el extremo inferior de la escala. La concentración del 95,3 % de la muestra en el nivel más bajo del pretest hace que esta última explicación sea especialmente verosímil. Los resultados se reportan aquí con esa advertencia explícita, y su interpretación se desarrolla en el apartado siguiente.

Discusión

Los datos muestran que, tras la aplicación del programa, el estudiantado se desplazó de forma masiva desde el nivel de desempeño más bajo hacia el umbral de competencia esperado, con una reducción simultánea de la variabilidad relativa del grupo. Este segundo hallazgo merece tanta atención como el primero. Una intervención que eleva la media pero mantiene o amplía la dispersión beneficia sobre todo a quienes ya partían de mejores condiciones; el descenso del coeficiente de variación de 21,7 % a 10,2 % sugiere lo contrario, es decir, que los estudiantes con puntuaciones iniciales más bajas fueron quienes más avanzaron. La correlación prácticamente nula entre ambas mediciones refuerza esa lectura, aunque también admite una explicación menos favorable, comentada más adelante entre las limitaciones. En términos de equidad educativa, ese es el resultado más valioso del estudio.

El patrón por dimensiones admite una lectura teórica coherente. El mayor avance se produjo en regularidad y equivalencia, la dimensión que exige identificar la estructura subyacente de un problema antes de aplicar un procedimiento. Ese es precisamente el tipo de discriminación que la práctica intercalada entrena, al impedir que el estudiante infiera la estrategia correcta a partir de la posición del ejercicio en la hoja de trabajo. El menor avance relativo en forma y movimiento resulta igualmente informativo: el razonamiento espacial se apoya en representaciones que, según la evidencia de neuroimagen, involucran patrones de activación parcialmente distintos de los del procesamiento simbólico (Koch et al., 2023), y es plausible que requiera un periodo de consolidación más extenso que el contemplado en el programa.

La convergencia con la literatura internacional es parcial y conviene precisarla. La dirección del efecto coincide con la evidencia acumulada sobre recuperación activa y práctica espaciada en contextos escolares (Carpenter et al., 2022; Latimier et al., 2021; Yang et al., 2021). Sin embargo, la magnitud observada se aparta considerablemente de la reportada por la síntesis metaanalítica más específica del dominio, que estimó un efecto pequeño a medio de la práctica espaciada en aprendizaje matemático y no halló evidencia concluyente de un efecto sistemático de la recuperación frente al reestudio (Murray et al., 2025). Esa discrepancia no invalida los resultados, pero sí obliga a una interpretación prudente: es más razonable atribuir la diferencia a las características del diseño y del punto de partida del grupo que a una eficacia superior del programa. Un antecedente útil en este sentido es el trabajo con estudiantes de secundaria que documentó mejoras en los puntajes junto con una reducción del exceso de confianza tras espaciar la práctica matemática (Emeny et al., 2021), un efecto de dimensiones mucho más modestas y obtenido bajo control experimental.

El componente de regulación emocional del error se apoyó en evidencia sólida respecto de la relación entre ansiedad matemática y desempeño (Barroso et al., 2021) y de la eficacia de las intervenciones dirigidas a esa relación (Codding et al., 2023; Sammallahti et al., 2023). El presente estudio, no obstante, no midió la ansiedad matemática, de modo que no puede afirmarse que la mejora observada haya operado por esa vía. Se trata de una limitación de diseño que conviene subsanar en trabajos futuros mediante la incorporación de una medida específica, lo que permitiría contrastar mecanismos en lugar de suponerlos.

Un segundo hallazgo con implicaciones prácticas se refiere al contexto tecnológico de la intervención. La evidencia reciente sobre descarga cognitiva indica que el uso frecuente de asistentes de inteligencia artificial se asocia con puntuaciones menores en pensamiento crítico, con una dependencia más marcada entre los usuarios más jóvenes (Gerlich, 2025). La recuperación activa opera como contrapeso estructural frente a ese fenómeno, porque exige la producción de la respuesta antes de que ninguna fuente externa pueda proporcionarla. La distinción relevante, siguiendo el marco de la cognición distribuida, no es entre usar o no usar la herramienta, sino entre delegar procesos periféricos y delegar aquellos que constituyen el objetivo mismo del aprendizaje (Grinschgl & Neubauer, 2022). Diseñar tareas que hagan explícita esa frontera parece hoy una tarea urgente de la didáctica de la matemática.

El estudio también permite formular una advertencia sobre el uso de la etiqueta neurodidáctica. La prevalencia de neuromitos entre docentes en ejercicio y en formación sigue siendo elevada (Novak-Geiger, 2023; Rato et al., 2022), y las propuestas que invocan el cerebro sin especificar qué evidencia sostiene cada componente contribuyen a esa persistencia. La buena noticia es que el problema resulta corregible: una formación breve y de bajo costo produce reducciones considerables y duraderas de esas creencias (Ruiz-Martín et al., 2022), y el trabajo con conceptos neurocientíficos modifica de manera sustantiva la manera en que los docentes conciben su propia práctica y a sus estudiantes (Chang et al., 2021). El desarrollo profesional situado, con acompañamiento en el aula y no únicamente con transmisión de contenidos, aparece como la modalidad con mejores perspectivas (Tan & Amiel, 2022). En coherencia con este planteamiento, el programa aquí descrito excluyó deliberadamente los componentes carentes de respaldo y documentó la base empírica de cada estrategia.

Limitaciones

La limitación principal es el diseño. Un estudio de un solo grupo con medidas pretest y postest no permite descartar explicaciones alternativas al efecto del programa. Entre las amenazas a la validez interna que resultan pertinentes en este caso figuran la maduración de los participantes a lo largo del periodo lectivo, el efecto de la doble aplicación del instrumento, la posible regresión a la media dada la extrema concentración de puntuaciones bajas en la medición inicial, y la historia, entendida como cualquier acontecimiento concurrente que haya podido incidir

en el aprendizaje matemático del grupo. La magnitud inusualmente alta del tamaño del efecto es, muy probablemente, expresión de estas amenazas antes que de una eficacia excepcional del programa.

En segundo lugar, el muestreo no probabilístico por conveniencia restringe la validez externa: los resultados describen a este grupo, en esta institución y en este periodo, y no autorizan inferencias sobre el bachillerato ecuatoriano en su conjunto. En tercer lugar, la ausencia de una medición de seguimiento impide determinar si las ganancias se mantienen en el tiempo, cuestión especialmente pertinente cuando la intervención se apoya en técnicas cuyo beneficio se define justamente por la retención a largo plazo. En cuarto lugar, no se midieron las variables mediadoras que el marco teórico postula, funciones ejecutivas, metacognición y ansiedad matemática, de modo que los mecanismos permanecen en el terreno de la hipótesis. Por último, el instrumento fue construido específicamente para este estudio, lo que exige prudencia en la comparación de sus resultados con los de investigaciones que emplean pruebas estandarizadas. A ello se añade un problema de sensibilidad: la concentración del 95,3 % de la muestra en el nivel más bajo del pretest indica que la prueba discriminó poco en el extremo inferior de la escala, y la correlación prácticamente nula entre ambas aplicaciones es compatible con esa falta de discriminación inicial. Un instrumento con mayor densidad de ítems en el rango bajo permitiría estimar el cambio con más precisión y reduciría el riesgo de sobrestimar el efecto.

Líneas futuras

La agenda que se desprende de estas limitaciones es relativamente clara. Un ensayo controlado con asignación aleatoria por conglomerados, con paralelos completos asignados a condición experimental o de control, resolvería la mayor parte de las amenazas señaladas y permitiría estimar el efecto con la precisión que hoy no se alcanza. La incorporación de medidas de funciones ejecutivas y de metacognición habilitaría análisis de mediación capaces de identificar por qué vía opera el programa (Muncer et al., 2022; Tette et al., 2026), y la evidencia disponible sobre entrenamiento de funciones ejecutivas en población escolar ofrece un punto de partida metodológico para ese diseño (Birtwistle et al., 2025). Una medición de seguimiento a los tres y a los seis meses permitiría evaluar la persistencia de las ganancias. Convendría, además, aplicar a este campo el estándar de escrutinio que otras áreas de la investigación educativa ya han

adoptado, tras constatar que intervenciones ampliamente difundidas ofrecían efectos considerablemente menores a los anunciados una vez controlada la calidad metodológica de los estudios primarios (Macnamara & Burgoyne, 2023). Finalmente, la discusión abierta sobre el equilibrio entre instrucción directa e indagación en el aprendizaje matemático (Sweller et al., 2024) sugiere que el propio contenido del programa merece ser sometido a comparación experimental entre variantes, y no solo contrastado con la enseñanza habitual.

Conclusiones

El programa de estrategias neurodidácticas se asocia con una mejora sustantiva del desempeño en competencias matemáticas de los estudiantes de tercero de bachillerato participantes. La proporción de estudiantes en el nivel Inicio descendió del 95,3 % al 1,6 %, la media aumentó 6,16 puntos sobre una escala de 20 y las diferencias entre ambos momentos resultaron estadísticamente significativas, $t(127) = 33{,}45$; $p < {,}001$; $d = 4{,}16$, con confirmación no paramétrica ($Z = 9{,}81$; $p < {,}001$; $r = {,}61$). La reducción del coeficiente de variación indica que la mejora alcanzó de manera más intensa a quienes partían de posiciones más desfavorables.

El avance fue transversal a las tres dimensiones evaluadas, con mayor intensidad en la resolución de problemas de regularidad y equivalencia y menor en la de forma y movimiento. El nivel Destacado permaneció marginal, lo que indica que el programa resulta eficaz para llevar al estudiantado hasta el umbral de competencia esperado, pero insuficiente, en la duración estudiada, para producir desempeños de excelencia.

Estas conclusiones deben leerse dentro de los límites del diseño empleado. Los datos son consistentes con la hipótesis de que el programa produjo la mejora, pero no la establecen de manera concluyente, y el tamaño del efecto observado excede lo que la literatura considera plausible para intervenciones de este tipo. Antes que un argumento a favor de la generalización inmediata de la propuesta, el estudio ofrece una base empírica documentada para el diseño de un ensayo controlado que permita estimar el efecto con las garantías que este no proporciona. Esa es, a juicio de los autores, la contribución más honesta que puede reclamarse para un trabajo de estas características.

Declaraciones

Agradecimientos. Los autores agradecen a las autoridades, al personal docente y al estudiantado de la institución participante por las facilidades brindadas para la ejecución del trabajo de campo.

Financiamiento. Investigación autofinanciada. Los autores no recibieron financiamiento externo para la realización de este estudio.

Conflicto de intereses. Los autores declaran no tener conflictos de intereses.

Contribución de autoría. **Elian Jamil Hernández Cueva:** Conceptualización, metodología, investigación, análisis formal y redacción del borrador original. **Evelyn Michelle Ordóñez Merino:** Curación de datos, investigación, metodología y redacción del borrador original. **Walter Darío Ramón Namcela:** Investigación, análisis formal, revisión y edición del manuscrito. **Jessica Rosalía Espinosa González:** Conceptualización, curación de datos, redacción, revisión y edición.

Todos los autores contribuyeron al desarrollo del estudio, participaron en la discusión e interpretación de los resultados, revisaron críticamente el manuscrito y aprobaron la versión final para su publicación.

Disponibilidad de datos. Los datos que sustentan los hallazgos de este estudio están disponibles previa solicitud razonable al autor de correspondencia.

Referencias

American Psychological Association. (2020). *Publication manual of the American Psychological Association* (7.ª ed.). https://doi.org/10.1037/0000165-000

Barroso, C., Ganley, C. M., McGraw, A. L., Geer, E. A., Hart, S. A., & Daucourt, M. C. (2021). A meta-analysis of the relation between math anxiety and math achievement. *Psychological Bulletin, 147*(2), 134-168. https://doi.org/10.1037/bul0000307

Birtwistle, E., Chernikova, O., Wünsch, M., & Niklas, F. (2025). Training of executive functions in children: A meta-analysis of cognitive training interventions. *SAGE Open, 15*(1). https://doi.org/10.1177/21582440241311060

Carpenter, S. K., Pan, S. C., & Butler, A. C. (2022). The science of effective learning with spacing and retrieval practice. *Nature Reviews Psychology, 1*(9), 496-511. https://doi.org/10.1038/s44159-022-00089-1

Chang, Z., Schwartz, M. S., Hinesley, V., & Dubinsky, J. M. (2021). Neuroscience concepts changed teachers' views of pedagogy and students. *Frontiers in Psychology, 12*, 685856. https://doi.org/10.3389/fpsyg.2021.685856

Codding, R. S., Goodridge, A. E., Hill, E., Kromminga, K. R., Chehayeb, R., Volpe, R. J., & Scheman, N. (2023). Meta-analysis of skill-based and therapeutic interventions to address math anxiety. *Journal of School Psychology, 100*, 101229. https://doi.org/10.1016/j.jsp.2023.101229

Emeny, W. G., Hartwig, M. K., & Rohrer, D. (2021). Spaced mathematics practice improves test scores and reduces overconfidence. *Applied Cognitive Psychology, 35*(4), 1082-1089. https://doi.org/10.1002/acp.3814

Gerlich, M. (2025). AI tools in society: Impacts on cognitive offloading and the future of critical thinking. *Societies, 15*(1), 6. https://doi.org/10.3390/soc15010006

Grinschgl, S., & Neubauer, A. C. (2022). Supporting cognition with modern technology: Distributed cognition today and in an AI-enhanced future. *Frontiers in Artificial Intelligence, 5*, 908261. https://doi.org/10.3389/frai.2022.908261

Iglesias-Sarmiento, V., Carriedo, N., Rodríguez-Villagra, O. A., & Pérez, L. (2023). Executive functioning skills and (low) math achievement in primary and secondary school. *Journal of Experimental Child Psychology, 235*, 105715. https://doi.org/10.1016/j.jecp.2023.105715

Kim, S. K., & Webb, S. (2022). The effects of spaced practice on second language learning: A meta-analysis. *Language Learning, 72*(1), 269-319. https://doi.org/10.1111/lang.12479

Koch, G. E., Libertus, M. E., Fiez, J. A., & Coutanche, M. N. (2023). Representations within the intraparietal sulcus distinguish numerical tasks and formats. *Journal of Cognitive Neuroscience, 35*(2), 226-240. https://doi.org/10.1162/jocn_a_01933

Latimier, A., Peyre, H., & Ramus, F. (2021). A meta-analytic review of the benefit of spacing out retrieval practice episodes on retention. *Educational Psychology Review, 33*(3), 959-987. https://doi.org/10.1007/s10648-020-09572-8

Leisman, G. (2023). Neuroscience in education: A bridge too far or one that has yet to be built. *Brain Sciences, 13*(1), 40. https://doi.org/10.3390/brainsci13010040

Macnamara, B. N., & Burgoyne, A. P. (2023). Do growth mindset interventions impact students' academic achievement? A systematic review and meta-analysis with recommendations for best practices. *Psychological Bulletin, 149*(3-4), 133-173. https://doi.org/10.1037/bul0000352

Muncer, G., Higham, P. A., Gosling, C. J., Cortese, S., Wood-Downie, H., & Hadwin, J. A. (2022). A meta-analysis investigating the association between metacognition and math performance in adolescence. *Educational Psychology Review, 34*(1), 301-334. https://doi.org/10.1007/s10648-021-09620-x

Murray, E., Horner, A. J., & Göbel, S. M. (2025). A meta-analytic review of the effectiveness of spacing and retrieval practice for mathematics learning. *Educational Psychology Review, 37*, 75. https://doi.org/10.1007/s10648-025-10035-1

Novak-Geiger, V. (2023). Prevalence of neuromyths among psychology students: Small differences to pre-service teachers. *Frontiers in Psychology, 14*, 1139911. https://doi.org/10.3389/fpsyg.2023.1139911

OCDE. (2023). *PISA 2022 results (Volume I): The state of learning and equity in education*. OECD Publishing. https://doi.org/10.1787/53f23881-en

Pan, S. C., & Carpenter, S. K. (2023). Prequestioning and pretesting effects: A review of empirical research, theoretical perspectives, and implications for educational practice. *Educational Psychology Review, 35*(4), 97. https://doi.org/10.1007/s10648-023-09814-5

Rato, J. R., Amorim, J., & Castro-Caldas, A. (2022). Looking for the brain inside the initial teacher training and outreach books in Portugal. *Frontiers in Psychology, 13*, 737136. https://doi.org/10.3389/fpsyg.2022.737136

Roell, M., Cachia, A., Matejko, A. A., Houdé, O., Ansari, D., & Borst, G. (2021). Sulcation of the intraparietal sulcus is related to symbolic but not non-symbolic number skills. *Developmental Cognitive Neuroscience, 51*, 100998. https://doi.org/10.1016/j.dcn.2021.100998

Ruiz-Martín, H., Portero-Tresserra, M., Martínez-Molina, A., & Ferrero, M. (2022). Tenacious educational neuromyths: Prevalence among teachers and an intervention. *Trends in Neuroscience and Education, 29*, 100192. https://doi.org/10.1016/j.tine.2022.100192

Sammallahti, E., Finell, J., Jonsson, B., & Korhonen, J. (2023). A meta-analysis of math anxiety interventions. *Journal of Numerical Cognition, 9*(2), 346-362. https://doi.org/10.5964/jnc.8401

Sweller, J. (2022). The role of evolutionary psychology in our understanding of human cognition: Consequences for cognitive load theory and instructional procedures. *Educational Psychology Review, 34*(4), 2229-2241. https://doi.org/10.1007/s10648-021-09647-0

Sweller, J. (2023). The development of cognitive load theory: Replication crises and incorporation of other theories can lead to theory expansion. *Educational Psychology Review, 35*(4), 95. https://doi.org/10.1007/s10648-023-09817-2

Sweller, J., Zhang, L., Ashman, G., Cobern, W., & Kirschner, P. A. (2024). Response to De Jong et al.'s (2023) paper "Let's talk evidence – The case for combining inquiry-based and direct

instruction". *Educational Research Review, 42*, 100584. https://doi.org/10.1016/j.edurev.2023.100584

Tan, Y. S. M., & Amiel, J. J. (2022). Teachers learning to apply neuroscience to classroom instruction: Case of professional development in British Columbia. *Professional Development in Education, 48*(1), 70-87. https://doi.org/10.1080/19415257.2019.1689522

Tette, P. P. M., Justi, C. N. G., & Justi, F. R. R. (2026). The relationship between executive functions and mathematics: A systematic review with meta-analysis of longitudinal studies. *Psicologia: Reflexão e Crítica, 39*. https://doi.org/10.1186/s41155-025-00362-1

UNESCO-OREALC & Laboratorio Latinoamericano de Evaluación de la Calidad de la Educación. (2021). *Estudio Regional Comparativo y Explicativo (ERCE 2019): Reporte nacional de resultados, Ecuador*. UNESCO. https://evaluaciones.evaluacion.gob.ec/archivosPD/uploads/dlm_uploads/2022/03/Estudio-Regional-Comparativo-y-Explicativo-ERCE-2019.pdf

Yang, C., Luo, L., Vadillo, M. A., Yu, R., & Shanks, D. R. (2021). Testing (quizzing) boosts classroom learning: A systematic and meta-analytic review. *Psychological Bulletin, 147*(4), 399-435. https://doi.org/10.1037/bul0000309