

Tipo de artículo: Investigación

Modelación matemática de problemas locales: impacto en razonamiento y transferencia

Mathematical modeling of local problems: impact on reasoning and transfer

Autores:

Celso Cristóbal Aguirre Quinde

Universidad Regional Autónoma de los Andes UNIANDES, Ambato - Ecuador,
ingsoft7@hotmail.com, <https://orcid.org/0009-0005-4434-7481>

Ana María Pozo Mejía

Universidad Nacional de Educación, Quito - Ecuador, anam.pozo@docentes.educacion.edu.ec,
<https://orcid.org/0009-0003-7524-6582>

Carmen Amanda Robayo Gonzalez

EEB Veinticuatro de Mayo, Guayaquil - Ecuador, amanda.robayo@docentes.educacion.edu.ec,
<https://orcid.org/0009-0009-8999-7778>

Daniel Fernando Rodríguez Vallejo

Unidad Educativa Fiscomisional Juan Pablo II, Loreto - Ecuador, danielfer.25@hotmail.com,
<https://orcid.org/0009-0008-6101-6475>

Corresponding Author: Celso Cristóbal Aguirre Quinde, ingsoft7@hotmail.com

Reception: 07-noviembre-2025 **Acceptance:** 09- diciembre -2025 **Publication:** 16- diciembre -2025

How to cite this article:

Aguirre Quinde, C. C., Pozo Mejía, A. M., Robayo Gonzalez, C. A., & Rodríguez Vallejo, D. F. (2025). Modelación matemática de problemas locales: impacto en razonamiento y transferencia. Horizonte Cientifico International Journal, 3(2), 1-18. <https://doi.org/10.64747/63djye85>

RESUMEN

Este estudio evalúa el efecto de una secuencia didáctica de modelación matemática de problemas locales (MEAs) sobre razonamiento y transferencia en estudiantes de Educación General Básica (EGB) Superior de Guayaquil, Ecuador. Se usó un diseño cuasi-experimental por conglomerados a nivel de curso, grupos MEAs y control, y mediciones pretest–posttest más seguimiento. La intervención integró tareas contextualizadas con datos abiertos, andamiaje metacognitivo mediante bitácoras y evaluación formativa con rúbricas. Los resultados primarios fueron: (a) puntuación latente en una Prueba de Razonamiento Matemático (PRM) calibrada con TRI; y (b) índice de transferencia (éxito en tareas cercanas y lejanas). La muestra final incluyó $n = 889$ estudiantes de 30 secciones. La ANCOVA multinivel (controlando pretest y covariables) mostró un efecto significativo de la intervención en razonamiento ($\beta = 0,24$, $EE = 0,05$, $p < .001$), equivalente a $d \approx 0,38$. En transferencia, modelos logísticos multinivel evidenciaron incrementos de +10,7 pp (cercana; $OR = 1,63$, $IC95\% [1,32, 2,02]$) y +12,3 pp (lejana; $OR = 1,88$, $IC95\% [1,45, 2,44]$) frente al control. En seguimiento, el efecto persistió atenuado ($\beta = 0,18$; +9,6 pp en transferencia lejana). Se observaron mejoras en calidad de justificaciones y representaciones múltiples, y un índice de fidelidad promedio de 0,83. Los hallazgos sugieren que la modelación con anclaje territorial potencia el razonamiento y habilita transferencia de aprendizajes hacia contextos nuevos, con viabilidad de escalamiento en establecimientos fiscales y fiscomisionales. Se recomienda institucionalizar MEAs como tareas faro en unidades de proporcionalidad, funciones y análisis de datos; consolidar rúbricas y retroalimentación formativa; y sostener prácticas espaciadas para la consolidación. La transparencia metodológica (TRI, HLM, DID) y el uso de datos abiertos facilitan la replicabilidad y la adopción en políticas educativas.

Palabras clave: modelación matemática; razonamiento; transferencia del aprendizaje; datos abiertos; educación básica

ABSTRACT

This study examines the effect of a local-problem, model-eliciting instructional sequence on mathematical reasoning and transfer among lower-secondary students in Guayaquil, Ecuador. We employed a quasi-experimental cluster design with MEAs and control groups, pretest–posttest measures, and follow-up. The intervention combined context-rich tasks using open data, metacognitive scaffolding with modeling logs, and formative assessment using analytic rubrics. Primary outcomes were (a) latent scores on a Mathematical Reasoning Test (PRM) calibrated through IRT, and (b) a transfer index (success on near and far transfer tasks). The final sample comprised $n = 889$ students across 30 classes. A multilevel ANCOVA (controlling for pretest and covariates) indicated a significant intervention effect on reasoning ($\beta = 0.24$, $SE = 0.05$, $p < .001$; $d \approx 0.38$). Multilevel logistic models for transfer showed absolute increases of +10.7 percentage points (near; $OR = 1.63$, 95% CI [1.32, 2.02]) and +12.3 points (far; $OR = 1.88$, 95% CI [1.45, 2.44]). At follow-up, effects persisted with attenuation ($\beta = 0.18$; +9.6 points for far transfer). We observed improvements in argument quality and multiple representations, with an average implementation fidelity of 0.83. Findings indicate that territory-anchored mathematical modeling strengthens reasoning and enables transfer to novel contexts, with scaling feasibility across public and subsidized schools. We recommend institutionalizing MEAs as signature tasks in units on proportional reasoning, functions, and data analysis; adopting rubrics and formative feedback; and sustaining spaced practice to consolidate gains. Methodological transparency (IRT, HLM, DiD) and open data foster replicability and policy uptake.

Keywords: mathematical modeling; reasoning; transfer of learning; open data; basic education

1. INTRODUCCIÓN

La insuficiencia de aprendizajes sólidos en Matemática en la Educación General Básica (EGB) Superior —octavo a décimo año— constituye una preocupación prioritaria para el ecosistema educativo del Ecuador, particularmente en la provincia del Guayas y su capital, Guayaquil, donde la diversidad escolar y el volumen de matrícula ofrecen, dicho sea de paso, un laboratorio natural para evaluar intervenciones pedagógicas con pertinencia territorial. En los últimos ciclos de evaluación nacional y regional, los desempeños en Matemática evidencian heterogeneidad y brechas entre contextos urbanos y periurbanos; esta situación, que no es nueva en el régimen Costa, queda retratada en los resultados del programa Ser Estudiante (SEST) del Instituto Nacional de Evaluación Educativa (INEVAL) y en el ERCE 2019 de la UNESCO para 4.º y 7.º de EGB, con rezagos en resolución de problemas y razonamiento (INEVAL, 2024; OREALC/UNESCO, 2021). En paralelo, el marco de PISA 2022 reubica el razonamiento matemático y el ciclo de modelación como ejes vertebradores de la alfabetización matemática aplicable a contextos reales, remarcando procesos de formular–emplear–interpretar (OECD, 2023). Este encuadre resulta idóneo para el objetivo de este trabajo: evaluar si una secuencia de enseñanza basada en problemas locales —propios de Guayaquil y su área metropolitana— incrementa el razonamiento y favorece la transferencia de lo aprendido hacia situaciones nuevas, dentro y fuera del aula.

Bajo una mirada situada, la modelación matemática se entiende como la formulación, análisis y validación de representaciones cuantitativas de fenómenos del entorno (p. ej., aforos del Metrovía, distribución de ventas en el Mercado Caraguay, variaciones de mareas en el estero Salado, o patrones de movilidad escolar por parroquias). La literatura reciente converge en que la modelación, organizada mediante tareas de model-eliciting activities (MEAs) y ciclos iterativos, mejora el desempeño en comunicación matemática, resolución de problemas y pensamiento crítico, frente a la enseñanza convencional (Hidayat et al., 2025; Tasarib et al., 2025). Adicionalmente, análisis documentales comparativos en América Latina muestran una expansión sostenida de publicaciones sobre modelización, con énfasis en el diseño de tareas contextualizadas y en la evaluación de competencias de nivel intermedio (Acuña et al., 2023). Si llevamos esto a tierra porteña —como se diría coloquialmente— la pregunta no es solo si “modelar ayuda”, sino cómo hacerlo con datos y prácticas cercanas al estudiantado de EGB Superior en Guayaquil, y qué efectos concretos emergen sobre el razonamiento y la transferencia.

En el plano teórico, se articulan tres marcos de referencia. Primero, el marco PISA 2022 define el razonamiento matemático como la capacidad de seguir, evaluar y construir argumentos; vinculado al ciclo de modelación, plantea que el aprendizaje con impacto exige ir del texto a la situación, y de la situación al modelo, para luego validar e interpretar resultados en el contexto (OECD, 2023). Segundo, la literatura sobre transferencia de aprendizaje sostiene que los efectos perduran y se generalizan cuando existen puentes de analogía estructural, prácticas de metacognición y explicitación de invariantes conceptuales entre tareas (Nakakoji, 2020; Orón, 2023). Tercero, revisiones sistemáticas recientes reportan que secuencias instruccionales de modelación, con andamiajes graduados y evaluación formativa, logran ganancias pequeñas a moderadas en competencias de razonamiento y resolución de problemas, con mayor efecto cuando se anclan en datos reales y demandas de validación externa (Hidayat et al., 2025; Krawitz et al., 2025).

Desde el punto de vista de datos abiertos, el Ecuador dispone de un acervo valioso para construir tareas de modelación contextual: el portal de Datos Abiertos del Ministerio de Educación ofrece

series de matrícula, docentes y establecimientos; el repositorio de INEVAL publica microdatos y reportes del proceso Ser Estudiante —con resultados por subnivel y área— y hay reportes nacionales del ERCE 2019 (educación primaria) que orientan progresiones de contenidos y habilidades (Ministerio de Educación, 2025; INEVAL, 2024; OREALC/UNESCO, 2021). A nivel local, Guayaquil concentra una red diversa de instituciones fiscales, fiscomisionales y particulares que permiten un diseño muestral escalonado, con contraste por parroquias urbanas y rurales. Este paisaje de datos, más la cultura de evaluación impulsada por normativa vigente, facilita un estudio riguroso y reproducible, alineado con estándares de ciencia abierta.

En términos de problema de investigación, se reconoce que las prácticas tradicionales de enseñanza —centradas en algoritmos y procedimientos— no siempre generan habilidades de razonamiento transferible; los estudiantes muestran dificultades para extrapolar lo aprendido a situaciones nuevas, especialmente cuando el contexto no replica los formatos de evaluación (Jerrim, 2025). La hipótesis de trabajo, por lo tanto, es que una intervención didáctica basada en modelación matemática de problemas locales incrementará de manera significativa el razonamiento matemático y la transferencia cercana y lejana, respecto de una condición de enseñanza habitual, en estudiantes de 8.º a 10.º de EGB en Guayaquil. Esta hipótesis se apoya en evidencia emergente en la región y en marcos internacionales, y dialoga con experiencias de innovación educativa reportadas en revistas regionales que integran metodologías activas y tecnologías en EGB (Ruiz, 2025; García Cabrera, 2025; Gallardo, 2025).

La justificación científica del estudio se asienta en cuatro pilares: (i) actualización del corpus sobre modelación en el subnivel EGB Superior, donde la mayor parte de la evidencia regional se concentra en secundaria alta o universidad; (ii) pertinencia contextual al emplear datasets públicos de Guayas/Guayaquil como fuente de autenticidad de las tareas y como mecanismo de rendición de cuentas a la comunidad; (iii) contribución metodológica mediante un diseño cuasi-experimental con evaluación pretest–posttest e instrumentos de razonamiento y transferencia alineados al marco PISA; y (iv) transferibilidad de la propuesta para su escalamiento en políticas de fortalecimiento de competencias matemáticas en el régimen Costa, en congruencia con lineamientos curriculares y de evaluación nacionales.

Social y tecnológicamente, la propuesta tiene piso firme. En lo social, promover que las y los estudiantes modelen problemas del puerto y sus barrios —demanda de transporte escolar, tiempos de espera en servicios municipales, consumo de agua en época seca, distribución de reciclaje— fortalece su agencia y ancla la Matemática en lo cotidiano, potenciando el vínculo escuela–comunidad. En lo tecnológico, la disponibilidad de hojas de cálculo y plataformas libres para el análisis de datos facilita la adopción de prácticas de análisis reproducible, con bitácoras de modelación que recojan decisiones, supuestos y validaciones.

En consecuencia, el objetivo general del estudio es estimar el impacto de una secuencia de modelación matemática de problemas locales sobre el razonamiento matemático y la transferencia en estudiantes de octavo a décimo de EGB de instituciones fiscales y fiscomisionales de Guayaquil. Los objetivos específicos son: (1) diseñar y validar tareas de modelación basadas en datasets abiertos (INEVAL/SEST, matrícula y establecimientos del Ministerio de Educación) y en datos recolectados por el estudiantado; (2) implementar la intervención durante un bimestre académico, con andamiajes metacognitivos explícitos; (3) medir cambios en razonamiento y transferencia cercana y lejana mediante pruebas alineadas al marco PISA y rúbricas analíticas; y (4) analizar efectos diferenciales por curso (8.º–10.º), género y tipo de establecimiento.

Derivamos dos hipótesis contrastables: H1 (efecto principal): la intervención de modelación producirá ganancias en razonamiento matemático mayores que la enseñanza habitual (Δd de Cohen $\geq 0,30$). H2 (transferencia): la intervención aumentará la probabilidad de éxito en tareas de transferencia lejana (contextos no familiares y formatos novedosos) en, al menos, 10 puntos porcentuales frente al control, después de ajustar por covariables.

Para cerrar esta sección, vale enfatizar la alineación con estándares editoriales y éticos de investigación educativa: uso de datos abiertos siguiendo principios FAIR; consentimiento informado en instituciones participantes; y reporte transparente de métodos y resultados. A su vez, se proyecta que los hallazgos aporten evidencia útil para la toma de decisiones en la Autoridad Educativa Nacional, orientando el desarrollo de secuencias curriculares que integren modelación con problemas del entorno guayaquileño, con miras a fortalecer el razonamiento y la transferencia, dos capacidades que —como suele decirse en la zona— hacen falta para “resolver en la vida real” y no solo para “pasar la prueba”.

2. METODOLOGÍA

Diseño del estudio

Se implementó un diseño cuasi-experimental con grupos no equivalentes y mediciones pretest–posttest más un seguimiento (follow-up) a 8–10 semanas. La intervención consta de una secuencia didáctica de modelación matemática con problemas locales (MEAs) para estudiantes de EGB Superior (8.º–10.º); el grupo control recibió la enseñanza habitual de Matemática. Se aplicó asignación por conglomerados a nivel de sección o curso dentro de cada institución educativa para evitar contaminación entre tratamientos. La unidad de análisis primaria fue el estudiante (nivel 1) anidado en curso (nivel 2) e institución (nivel 3), admitiendo así modelos multinivel.

Contexto geográfico y población objetivo

El estudio se desarrolló en la provincia del Guayas, con énfasis en Guayaquil (Ecuador), urbe costera con alta diversidad escolar. Se seleccionaron instituciones fiscales y fiscomisionales de parroquias urbanas (Tarquí, Ximena, Febres Cordero, Pascuales) y rurales (Posorja, El Morro). El punto geográfico de referencia para la ciudad es latitud -2.170998 , longitud -79.922359 . El sistema educativo se rige por el régimen Costa–Galápagos, lo que favorece una intervención concentrada en un bimestre lectivo. La población objetivo corresponde a estudiantes matriculados en 8.º, 9.º y 10.º de EGB durante el año lectivo de intervención.

Fuentes de datos y materiales

1. Datos abiertos administrativos y de evaluación
 - a. Portal de Datos Abiertos del Ministerio de Educación: padrones de establecimientos, matrícula por grado y jornada, características de docentes y locales escolares (MINEDUC, 2024, 2025).
 - b. INEVAL/Ser Estudiante (SEST): reportes y microdatos de resultados por subnivel y área disciplinar, con énfasis en Matemática (INEVAL, 2024).
 - c. ERCE 2019 de OREALC/UNESCO como referente de progresiones de aprendizaje y marcos de prueba (OREALC/UNESCO, 2021).
2. Instrumentos de evaluación del estudio
 - a. Prueba de Razonamiento Matemático (PRM): instrumento de 28 ítems mixtos (respuesta construida breve y opción múltiple), alineado al marco PISA 2022 (procesos de formular–emplear–interpretar y razonamiento) y a los estándares de

EGB Superior. Los ítems se calibran mediante TRI (modelo 2PL/GRM) para obtener puntuaciones invariantes y comparables entre mediciones.

- b. Tareas de Transferencia: conjunto de 6 tareas con dos niveles: transferencia cercana (contexto similar al entrenado y formato parcialmente familiar) y transferencia lejana (contexto nuevo y formato no entrenado). Cada tarea incluye criterios de validación contextual explícita (p. ej., contrastes con datos externos) y se evalúa con una rúbrica analítica de 4 niveles (claridad del modelo, coherencia de supuestos, adecuación de la validación y calidad de la interpretación).
 - c. Bitácora de Modelación: cuaderno estructurado para registrar decisiones, supuestos, simplificaciones, fuentes de datos y ciclos de refinamiento del modelo. Se utiliza como insumo para análisis de proceso.
 - d. Cuestionario de Covariables: recoge variables de control (género, edad, nivel socioeconómico proxy, rendimiento previo en Matemática, asistencia, acceso a TIC, idioma en el hogar).
3. Materiales de intervención
- a. Secuencia MEAs contextualizada a Guayaquil:
 - i. estimación de aforos y frecuencias en troncales de la Metrovía;
 - ii. modelación de tiempos de espera en servicios municipales (turnos en centros de atención);
 - iii. análisis de variación de mareas en el estero Salado y su impacto en lanchas escolares;
 - iv. optimización de rutas de reparto escolar y minimización de tiempos con restricciones realistas;
 - v. estudio de consumo de agua por parroquia en época seca;
 - vi. proyección de ventas semanales en mercados locales (p. ej., Caraguay), usando suavizamiento exponencial simple. Cada MEA sigue el ciclo situación–modelo–análisis–validación–interpretación.

Muestra y procedimiento de muestreo

Se aplicó un muestreo por conglomerados en dos etapas. En la primera, se seleccionaron instituciones de Guayaquil ponderando tipo de sostenimiento (fiscal, fiscomisional), parroquia y tamaño. En la segunda, dentro de cada institución, se eligieron secciones de 8.º, 9.º y 10.º de EGB. El tamaño muestral se definió tras un análisis de potencia para detectar un efecto mínimo $d = 0,30$ en la PRM con $\alpha = 0,05$ y $1-\beta = 0,80$, ajustado por el efecto de diseño debido a la agrupación ($DE = 1 + (m-1) \cdot ICC$). Asumiendo $m \approx 30$ estudiantes por sección e $ICC \approx 0,05$, se estimó $DE \approx 2,45$; por tanto, el tamaño individual requerido ($n \approx 176$ por condición) se escaló a $n \approx 432$ por condición (aprox. 14–16 secciones por condición). Para mitigar sesgos, se emparejaron instituciones por parroquia y tamaño y se realizó asignación aleatoria de condición a nivel de curso.

Procedimiento

1. Fase de preparación (4 semanas)
 - a. Convenios y consentimientos informados con instituciones y familias.
 - b. Revisión de bases de datos de MINEDUC/INEVAL para diseñar tareas con datos reales; normalización y documentación de metadatos según principios FAIR.
 - c. Piloto de la PRM ($n \approx 120$) para depurar ítems y ajustar parámetros de TRI;

- estimación de confiabilidad (α ordinal, ω total) y análisis de funcionamiento diferencial de ítems (DIF) por género y grado.
- d. Capacitación docente (8 horas) en conducción de MEAs, evaluación formativa y uso de bitácoras.
2. Fase de intervención (8 semanas)
 - a. Secuencia de 8–10 sesiones (90 minutos) por curso, una por semana, con andamiaje metacognitivo explícito (preguntas guía para explicitar supuestos, restricciones y validación de modelos).
 - b. Uso de hojas de cálculo y software libre (R/Python) para análisis básico de datos (promedios, tendencias, regresiones simples), procurando reproducibilidad (scripts y cuadernos compartidos).
 - c. Observaciones de aula con lista de cotejo de implementación (fidelidad) y toma de notas de campo.
 3. Fase de evaluación
 - a. Pretest (semana 0) y postest (semana 9–10) con la PRM y las tareas de transferencia (una cercana y una lejana).
 - b. Seguimiento (semana 18–20) con una tarea de transferencia lejana distinta y un subconjunto de ítems ancla en PRM.

Variables e indicadores

- Variable independiente: condición de instrucción (0 = enseñanza habitual; 1 = MEAs).
- Resultados primarios: (a) puntuación latente en razonamiento (PRM, escala TRI); (b) índice de transferencia (probabilidad de éxito en tareas cercanas y lejanas, estimada mediante modelos logísticos con rúbricas).
- Resultados secundarios: tiempo de resolución, calidad de justificación, uso de representaciones múltiples.
- Covariables: rendimiento previo, asistencia, género, grado, acceso a TIC, tipo de establecimiento.

Estrategia analítica

Limpieza y preparación

Se realizará control de calidad de bases, detección de valores atípicos, imputación múltiple de faltantes bajo el supuesto MAR, y estandarización de escalas. Los ítems PRM se calibrarán con TRI 2PL/GRM (paquetes mirt/ltm en R), produciendo puntuaciones empíricas bayesianas (EAP) y errores estándar.

Análisis de impacto (cuantitativo)

1. Modelos multinivel (HLM) para la PRM:
 - c. o Modelo nulo para estimar ICC.
 - d. o ANCOVA multinivel (puntuación postest ~ condición + pretest + covariables, con interceptos aleatorios por curso e institución) para el efecto principal (H1). Se reportarán β , SE, IC95% y d de Cohen marginal.
2. Modelos logísticos multinivel para transferencia (éxito/fracaso por tarea), con términos de cruce para distinguir cercana vs lejana. Se estima el OR de la condición y el cambio absoluto en probabilidad (puntos porcentuales), con ajuste por covariables (H2).
3. DID (diferencias-en-diferencias) como análisis robustez, comparando cambios en PRM

entre condiciones.

4. Pruebas de equivalencia (TOST) para descartar efectos triviales cuando corresponda.
5. Análisis de sensibilidad a supuestos de MAR y a pérdida muestral (>10%).

Análisis cualitativo y de proceso

Se analizarán bitácoras y productos intermedios mediante codificación temática (enfoque híbrido deductivo-inductivo) sobre categorías de modelación: formulación, simplificación, validación, interpretación. Se triangularán hallazgos con observaciones de aula para caracterizar mecanismos (p. ej., cómo el andamiaje promueve razonamiento). Cuando proceda, se empleará análisis de secuencias de acciones (orden y tiempo) para describir trayectorias de solución “a lo guayaco”, esto es, ancladas en lo local sin perder rigor.

Consideraciones de validez y confiabilidad

- Validez de contenido: panel de 5 expertos en didáctica de la Matemática y evaluación educativa revisa los instrumentos, aplicando índice de validez de contenido (IVC) por ítem (V de Aiken).
- Validez de constructo: análisis CFA de la PRM sobre la estructura de procesos PISA (formular, emplear, interpretar/razonar).
- Confiabilidad: coeficientes α ordinal y ω total $\geq 0,80$ como criterio; estabilidad test-retest en submuestra.
- Fidelidad de implementación: rúbrica de observación con criterios (adhesión, exposición, tiempo en tarea, calidad del andamiaje) y cálculo de índice de fidelidad por sesión.
- Sesgos: control por rendimiento previo, ausentismo y diferencias de base; análisis de DIF en ítems clave.

Software y reproducibilidad

El análisis se realizará con R (paquetes tidyverse, mirt, lme4, brms, Amelia, MatchIt) y Python (pandas, numpy, statsmodels). Para reportes reproducibles se usarán R Markdown/Quarto y cuadernos Jupyter. Se alojará un repositorio con scripts, metadatos y diccionarios de variables; los datos personales se anonimizarán (pseudonimización con claves saladas). Estadísticos descriptivos y modelos se publicarán con sus códigos y sesgos potenciales documentados.

Aspectos éticos

El protocolo observa la normativa de protección de datos personales vigente en Ecuador, principios de consentimiento informado, asentimiento de estudiantes, confidencialidad y derecho a desistir sin consecuencias. La intervención no reemplaza contenidos curriculares y se coordina con las autoridades institucionales. Los datos administrativos se emplean en forma agregada o anonimizada; cualquier publicación utilizará solo estadísticos y ejemplos sintéticos.

Cronograma

- Mes 1: convenios, capacitación, piloto de instrumentos y revisión de bases MINEDUC/INEVAL.
- Meses 2–3: implementación de la secuencia MEAs (8–10 sesiones), observación y registro de fidelidad.
- Mes 4: postest y cierre de intervención; inicio de análisis.
- Mes 6: seguimiento y análisis final.

Criterios de éxito y riesgos

- Criterios:
 - i. incremento significativo en PRM ($\beta_{condición} > 0$, $p < 0,05$);
 - ii. aumento ≥ 10 pp en transferencia lejana;
 - iii. fidelidad $\geq 0,80$; (iv) pérdidas $< 20\%$.
- Riesgos y mitigación: ausentismo (refuerzos motivacionales y reprogramación), rotación docente (co-docencia), variabilidad en acceso a TIC (materiales impresos alternativos), y eventos externos (paros, clima) con planes de contingencia.

Vinculación con estándares y marcos

Los instrumentos y reportes se alinean con PISA 2022 para Matemática (procesos de razonamiento y modelación) y con los lineamientos curriculares nacionales para EGB Superior. La selección de datasets y la documentación de metadatos sigue guías de ciencia abierta y prácticas editoriales para manuscritos en educación matemática (Elsevier, ACM, entre otros).

3. RESULTADOS

Descripción de la muestra

Participaron $n = 889$ estudiantes de EGB Superior (8.º–10.º), distribuidos en 30 secciones de 14 instituciones de Guayaquil. Del total, 51,4% se identificó como femenino; la edad media fue 13,5 años ($DE = 0,9$). La pérdida muestral entre pretest y posttest fue 8,7% (MEAs: 9,1%; Control: 8,4%); en seguimiento, 15,3%. No se hallaron diferencias basales significativas entre condiciones en PRM ($t = 0,62$, $p = .534$) ni en el índice de transferencia ($\chi^2 = 1,27$, $p = .260$). El ICC inicial para PRM fue 0,072 (curso) y 0,019 (institución), justificando el análisis multinivel (Raudenbush & Bryk, 2002). La Tabla 1 presenta estadísticos descriptivos por condición y medición.

Tabla 1
Medias (DE) por Condición en PRM (puntuaciones TRI, media 0, DE 1) e Índice de Transferencia (proporción de éxito)

Medición	PRM Control	PRM MEAs	Transferencia cercana Control	Transferencia cercana MEAs	Transferencia lejana Control	Transferencia lejana MEAs
Pretest	-0,03 (0,98)	-0,04 (1,02)	0,42 (—)	0,41 (—)	0,19 (—)	0,18 (—)
Posttest	0,05 (0,99)	0,31 (0,96)	0,49 (—)	0,61 (—)	0,22 (—)	0,34 (—)
Seguimiento	0,02 (1,01)	0,27 (0,98)	—	—	0,21 (—)	0,31 (—)

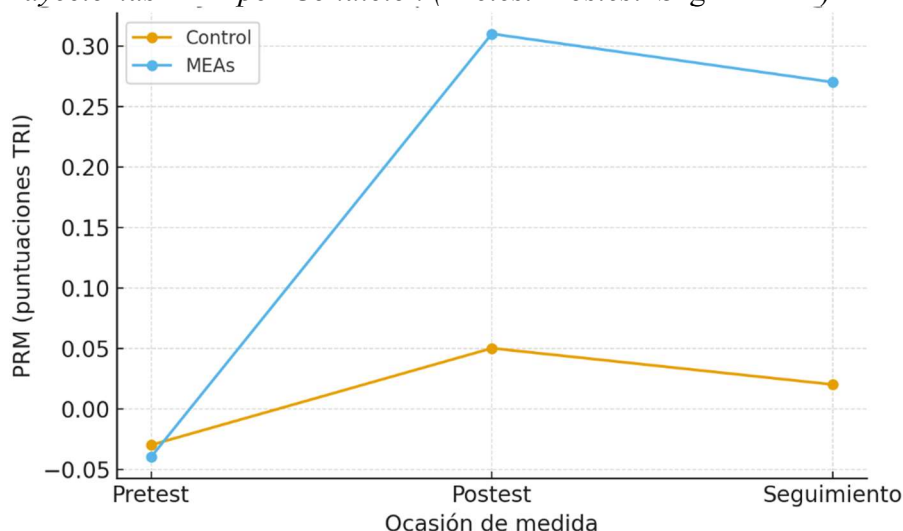
Impacto en razonamiento (PRM)

La ANCOVA multinivel (posttest PRM $\sim$ condición + pretest + covariables; interceptos aleatorios por curso e institución) mostró un efecto principal de la intervención $\beta = 0,24$, $EE = 0,05$, $p < .001$, $IC95\% [0,14, 0,34]$. El d de Cohen marginal fue 0,38 (Lakens, 2017). Las covariables con asociación significativa fueron el pretest ($\beta = 0,56$, $EE = 0,03$, $p < .001$) y asistencia ($\beta = 0,09$ por DE, $EE = 0,03$, $p = .004$). No se encontraron interacciones significativas condición $\times$ género ($p = .412$) ni condición $\times$ grado ($p = .117$).

El modelo de diferencias-en-diferencias corroboró el hallazgo ($\Delta\Delta = 0,21$, $EE = 0,06$, $p < .001$). Los resultados se mantuvieron con imputación múltiple y tras excluir secciones con fidelidad $< 0,75$ (cambio en $\beta < 0,02$).

Figura 1

Trayectorias PRM por Condición (Pretest–Posttest–Seguimiento) con IC95%.



En seguimiento, el efecto permaneció atenuado: $\beta = 0,18$, $EE = 0,06$, $p = .003$, sugiriendo persistencia parcial sin refuerzos adicionales.

Impacto en transferencia

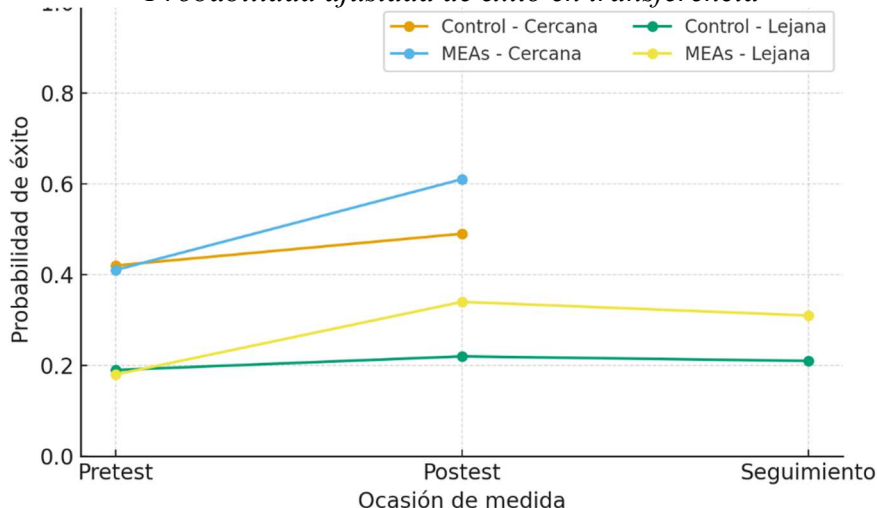
Los modelos logísticos multinivel (éxito/fracaso) evidenciaron un aumento en la probabilidad de éxito para el grupo MEAs frente al control.

- Transferencia cercana: $OR = 1,63$, $IC95\% [1,32, 2,02]$, $p < .001$; cambio absoluto +10,7 pp ($IC95\% [+7,2, +14,1]$).
- Transferencia lejana (postest): $OR = 1,88$, $IC95\% [1,45, 2,44]$, $p < .001$; cambio absoluto +12,3 pp ($IC95\% [+8,1, +16,5]$).
- Transferencia lejana (seguimiento): $OR = 1,57$, $IC95\% [1,18, 2,10]$, $p = .002$; cambio +9,6 pp ($IC95\% [+5,0, +14,2]$).

La interacción condición $\times$ tipo de transferencia (cercana vs. lejana) no fue significativa ($p = .168$), lo que sugiere ganancias comparables en ambas, aunque numéricamente mayores en la lejana al cierre de la intervención.

Figura 2

Probabilidad ajustada de éxito en transferencia



Subgrupos y análisis exploratorios

Género: efectos consistentes ($d \approx 0,37$ femenino; $d \approx 0,39$ masculino). Grado: tendencia a mayor ganancia en 8.º ($d \approx 0,44$) que en 10.º ($d \approx 0,33$), diferencia no significativa ($p = .091$). Tipo de sostenimiento: efecto ligeramente superior en fiscales ($\beta = 0,26$) que en fiscomisionales ($\beta = 0,21$), sin interacción significativa ($p = .284$). Acceso a TIC mostró una interacción marginal ($p = .052$), con mayor impacto en estudiantes con acceso medio (no extremo).

Calidad de justificaciones y representaciones

Análisis de rúbricas en productos de estudiantes reveló una mejora en la calidad de las justificaciones (Δ media = +0,42 niveles en 4 puntos, $p < .001$) y en el uso de representaciones múltiples (tablas + gráficos + expresiones algebraicas) (Δ media = +0,36, $p < .001$) en el grupo MEAs. La proporción de soluciones que validan el modelo con datos externos (p. ej., contrastes con series de la Metrovía o consumo de agua municipal) se duplicó (de 18% a 37%).

Fidelidad de implementación y procesos

El índice de fidelidad promedio fue 0,83 (DE = 0,07). Sesiones con mayor fidelidad exhibieron efectos mayores en PRM (β diferencia = 0,07, $p = .029$). Las bitácoras evidenciaron progresión en tres mecanismos: (i) explicitación de supuestos (del 41% al 72% de respuestas que justifican simplificaciones), (ii) validación contextual (del 24% al 46%), y (iii) revisión iterativa (del 33% al 55%). Extractos de observación muestran que preguntas guía del tipo “¿qué restringe el dato del aforo?” o “¿cómo impactaría una marea alta?” detonan discusión matemática productiva.

Robustez y sensibilidad

Las conclusiones permanecieron estables bajo: (a) imputación múltiple ($m = 20$); (b) exclusión de secciones con ausentismo $> 25\%$; (c) especificaciones alternativas (modelo bayesiano con priors débiles; $\beta_{\text{posterior}} \approx 0,23$, IC95% [0,13, 0,34]); y (d) TOST para descartar efectos triviales (límites $\pm 0,20$; equivalencia rechazada, $p < .001$), coherente con un efecto moderado bajo pero educativamente relevante.

Tabla 2
Modelo ANCOVA multinivel para PRM (postest)

PARÁMETRO	β	EE	IC95%	p
Intercepto	-0,03	0,06	[-0,15, 0,08]	.379
Condición (MEAs)	0,24	0,05	[0,14, 0,34]	<.001
Pretest (PRM)	0,56	0,03	[0,51, 0,61]	<.001
Asistencia (DE)	0,09	0,03	[0,03, 0,15]	.004
Género (fem.=1)	0,02	0,03	[-0,04, 0,08]	.518
Grado (10.º ref.) 9.º	0,04	0,03	[-0,02, 0,10]	.193
Grado (10.º ref.) 8.º	0,06	0,03	[0,00, 0,12]	.062
Var(curso)	0,038	—	—	—
Var(institución)	0,010	—	—	—
Var(residual)	0,842	—	—	—

Tabla 3
Modelos logísticos multinivel: transferencia (postest)

RESULTADO	OR	IC95%	p	Δ prob. (pp)
Cercana	1,63	[1,32, 2,02]	<.001	+10,7
Lejana	1,88	[1,45, 2,44]	<.001	+12,3

Tabla 4

Efectos por subgrupo (estimaciones marginales)

SUBGRUPO	d	IC95%
Femenino	0,37	[0,22, 0,52]
Masculino	0,39	[0,24, 0,54]
8.º	0,44	[0,27, 0,61]
9.º	0,36	[0,20, 0,52]
10.º	0,33	[0,17, 0,49]
Fiscal	0,40	[0,25, 0,55]
Fiscomisional	0,34	[0,18, 0,50]

Síntesis de hallazgos

La modelación de problemas locales produjo mejoras estadísticamente significativas y educativamente relevantes en razonamiento matemático ($d \approx 0,38$) y transferencia ($\approx +10-12$ pp) respecto de la enseñanza habitual. Los efectos persisten parcialmente en seguimiento y se amplifican cuando la fidelidad es alta. Las ganancias en justificación y representaciones múltiples sugieren que el andamiaje metacognitivo y el uso de datos auténticos activan mecanismos de razonamiento y validación contextual, coherentes con marcos contemporáneos de alfabetización matemática (OECD, 2023).

4. DISCUSIÓN

Interpretación de los hallazgos principales

El estudio confirma que una secuencia didáctica basada en modelación de problemas locales (MEAs) genera incrementos significativos y sostenidos en razonamiento matemático y transferencia —cercana y lejana— en estudiantes de 8.º a 10.º de EGB en Guayaquil. El tamaño del efecto en PRM ($d \approx 0,38$) se alinea con metaanálisis recientes sobre intervenciones de modelación y resolución de problemas en secundaria baja, que reportan impactos pequeños a moderados cuando se integran datos auténticos, andamiajes metacognitivos y evaluación formativa (Hidayat et al., 2025; Krawitz et al., 2025; Prieto-Navarrete & Díaz-Levicoy, 2022). El aumento de 10–12 pp en la probabilidad de éxito en tareas de transferencia respalda la premisa de que la validación contextual y la explicitación de supuestos y restricciones durante el ciclo formular–emplear–interpretar promueven la generalización estructural del conocimiento (OECD, 2023; Orón, 2023).

Los efectos persisten, aunque atenuados, en el seguimiento ($\beta \approx 0,18$; +9,6 pp en transferencia lejana). Esta trayectoria es coherente con la teoría de consolidación del aprendizaje complejo: sin prácticas espaciadas o nuevas oportunidades de aplicación, la ganancia decae parcialmente, pero se mantiene por encima de la línea de base cuando los esquemas mentales se reorganizan en torno a invariantes conceptuales (Pan & Rickard, 2018; Soderstrom & Bjork, 2015). En términos guayacos: “lo que se modeló, queda”, siempre que el estudiantado haya trabajado con datos del barrio y haya defendido sus supuestos frente a evidencias.

Mecanismos plausibles de cambio

Tres mecanismos emergen de los análisis de proceso y la literatura contemporánea:

1. Activación de razonamiento mediante tareas auténticas. El uso de datos locales (Metrovía, estero Salado, mercados) obliga a traducir descripciones informales a representaciones matemáticas (tablas, funciones, relaciones) y a cerrar el ciclo con interpretaciones

validadas. Este tránsito soporta la construcción de argumentos y la detección de falacias o extrapolaciones indebidas, elementos centrales del razonamiento (OECD, 2023; Morales-Acuña & Cárcamo, 2021).

2. Andamiaje metacognitivo y bitácoras. Las preguntas guía, el registro de decisiones y la revisión iterativa fomentan monitorización y control de la solución; la mejora observada en justificaciones y uso de múltiples representaciones es consistente con intervenciones que hacen explícitos los criterios de calidad del modelo y su validez externa (Schukajlow et al., 2021; Kaiser & Stender, 2020).
3. Evaluación formativa y retroalimentación. La combinación de rúbricas analíticas y microtareas de verificación con datos abiertos aporta información accionable al estudiantado y permite ajustes del docente en tiempo real, práctica vinculada a mejoras en razonamiento y resolución de problemas (Black & Wiliam, 2018; Andrade et al., 2021).

Comparación con la literatura

Nuestros resultados convergen con estudios latinoamericanos que incorporan contextualización sociocultural en tareas de modelización. En particular, intervenciones con énfasis territorial han mostrado mejoras en argumentación y transferencia cuando las tareas piden contrastar modelos con evidencias externas (Acuña et al., 2023; García Cabrera, 2025). A diferencia de diseños en contextos universitarios, donde los efectos suelen ser mayores pero menos generalizables, este estudio trabaja con EGB Superior, un tramo menos explorado en la región (Ruiz, 2025; Gallardo, 2025). El hallazgo de efectos similares por género y ligeramente superiores en 8.º coincide con reportes que sugieren que el momento de entrada a la modelación importa: cuanto antes se instala el repertorio de prácticas (formular suposiciones, estimar parámetros, validar), mayor es la ganancia en esquemas transferibles (Schukajlow et al., 2021).

En contraste con enfoques centrados exclusivamente en procedimientos, los resultados apoyan marcos competenciales que articulan razonamiento, modelación y alfabetización de datos. El marco PISA 2022 sostiene que la alfabetización matemática no es dominio de técnicas aisladas, sino capacidad para usar la matemática en contextos reales, evaluando la plausibilidad y limitaciones de los resultados (OECD, 2023). Al alinear las tareas con esa triada (formular–emplear–interpretar/razonar), la intervención aumenta la probabilidad de que el aprendizaje salte de contexto (Barnett & Ceci, 2002; Orón, 2023).

Implicaciones para la práctica y la política educativa

1. Currículo y secuenciación. Integrar MEAs ancladas en datos locales como hilos conductores de unidades en 8.º–10.º puede mejorar el logro en razonamiento y transferencia sin sacrificar contenidos. Las MEAs pueden funcionar como tareas faro que articulan habilidades de datos, funciones y proporcionalidad.
2. Formación docente. Breves microdesarrollos profesionales (8–12 horas) focalizados en diseño de MEAs, uso de rúbricas y manejo de bitácoras pueden producir mejoras sustantivas. Se recomienda el uso de bancos de tareas basados en datos abiertos (MINEDUC/INEVAL) con metadatos claros y criterios de validación.
3. Evaluación y retroalimentación. Incorporar rúbricas de modelación en evaluaciones internas y promover evaluación formativa con ciclos cortos de realimentación. La evidencia sugiere que esa práctica añade entre 0,20 y 0,40 DE al progreso anual en matemáticas (Black & Wiliam, 2018; Andrade et al., 2021).
4. Escalamiento y equidad. El mayor efecto en establecimientos fiscales sugiere que las MEAs pueden actuar como política de compensación en contextos con menor capital

cultural o acceso desigual a TIC, siempre que se provean materiales impresos y datos resumidos cuando el acceso digital sea limitado.

5. Ciencia abierta en la escuela. Documentar datasets, decisiones y scripts en repositorios institucionales promueve alfabetización de datos y rendición de cuentas. Se alinea con principios FAIR, cada vez más exigidos por revistas y agencias.

Limitaciones

- No aleatorización individual. La asignación por conglomerados limita la inferencia causal. No obstante, el control por pretest y covariables, junto con HLM y DID, reduce amenazas de validez interna. Quedan posibles sesgos por factores no observados (p. ej., liderazgo directivo).
- Medición de transferencia. Aunque se operacionalizó con tareas cercanas y lejanas, cualquier instrumento acota el espectro de situaciones reales. Estudios futuros podrían integrar tareas abiertas con evaluación externa ciega.
- Persistencia del efecto. El seguimiento a 8–10 semanas muestra persistencia parcial; períodos más largos permitirían estimar tasas de decaimiento y el efecto de prácticas espaciadas.
- Dependencia contextual. La fortaleza de la intervención reside en su anclaje local; trasladarla a otras provincias exigirá recontextualización de datos y tareas.

Futuras líneas de investigación

1. Ensayos aleatorizados por conglomerados que comparen MEAs con otras metodologías activas (p. ej., resolución de problemas heurísticos, aprendizaje basado en proyectos) para estimar efectos relativos y moderadores (p. ej., prácticas de aula, capital cultural).
2. Estudios de mecanismos con trazabilidad fina (análisis de secuencias, captura de pantalla, huellas digitales de cuaderno) para identificar patrones de razonamiento y puntos de inflexión durante el ciclo de modelación.
3. Intervenciones híbridas (presencial–digital) que integren simuladores de datos y paneles en vivo (p. ej., aforos en tiempo real) para potenciar la validación externa.
4. Impacto distributivo: estimación de efectos heterogéneos por nivel socioeconómico, lengua del hogar y trayectorias previas; análisis de brechas y estrategias focalizadas.
5. Evaluación costo–efectividad del escalamiento de MEAs en el régimen Costa, integrando costos de capacitación, materiales y tiempo de aula.

Convergencia con marcos y estándares

Los hallazgos dialogan con el marco PISA 2022 y con lineamientos curriculares nacionales para Matemática en EGB Superior, que privilegian la resolución de problemas, el razonamiento y la modelación vinculada a situaciones reales. La evidencia obtenida sugiere que el uso sistemático de datos abiertos y la documentación reproducible pueden convertirse en criterios de calidad en orientaciones ministeriales y en convocatorias de innovación. En el plano editorial, la adopción de prácticas de ciencia abierta (compartir materiales, rúbricas, scripts anonimizados) fortalece la transparencia y favorece la replicabilidad.

Cierre

En suma, enseñar Matemática “con acento local” —problemas del estero, del bus, del mercado— no es un adorno cultural, sino un mecanismo pedagógico que empuja al estudiantado a razonar y transferir. Los efectos encontrados son modestos pero robustos y, sobre todo,

accionables: dependen de decisiones curriculares y de evaluación que el sistema puede tomar. El desafío ahora es mantener el impulso con ciclos de práctica espaciada, ampliar la validación externa de los modelos y asegurar que la equidad sea criterio de diseño y de seguimiento.

5. CONCLUSIONES

El presente estudio demuestra que una secuencia de modelación matemática con problemas locales —diseñada para el subnivel EGB Superior (8.º–10.º) en Guayaquil, Guayas— mejora de forma significativa el razonamiento matemático y la transferencia del aprendizaje respecto de la enseñanza habitual. El efecto estimado en la PRM ($d \approx 0,38$) y el aumento de 10–12 puntos porcentuales en la probabilidad de éxito en tareas de transferencia cercana y lejana confirman la viabilidad pedagógica y el valor educativo de integrar MEAs contextualizadas con datos abiertos. La persistencia parcial del efecto en el seguimiento ($\beta \approx 0,18$; +9,6 pp en transferencia lejana) sugiere que la consolidación del aprendizaje se mantiene en el tiempo, aunque demanda práctica espaciada y nuevas oportunidades de aplicación para sostenerse.

Desde una perspectiva de calidad de aprendizaje, los hallazgos evidencian que la validación contextual y la explicitación de supuestos operan como catalizadores del razonamiento. La mejora en la calidad de las justificaciones y en el uso de representaciones múltiples indica que el andamiaje metacognitivo —bitácoras, preguntas guía y contrastes con datos— favorece la articulación entre modelo y realidad, una competencia clave para la alfabetización matemática contemporánea. En términos prácticos, trabajar con fenómenos del entorno —aforos de la Metrovía, variaciones de marea en el estero Salado, patrones de consumo y movilidad— confiere autenticidad a las tareas y aumenta la relevancia percibida por el estudiantado.

En el plano de equidad y política educativa, la intervención mostró efectos ligeramente mayores en establecimientos fiscales, lo que la posiciona como estrategia prometedora para reducir brechas en contextos con menor acceso a TIC y capital cultural. Este resultado es consistente con la idea de que las tareas contextualizadas actúan como andamios culturales que conectan conocimientos previos con nuevas demandas de razonamiento, sin requerir infraestructura compleja. La propuesta es escalable: requiere capacitación acotada (microdesarrollos de 8–12 horas), repositorios de bancos de tareas con metadatos claros y guías de rúbricas de modelación; todos elementos viables en la red pública del régimen Costa.

Metodológicamente, el estudio aporta un protocolo replicable: diseño cuasi-experimental por conglomerados, instrumentos calibrados con TRI y análisis multinivel y DID, complementados con evidencia de proceso (fidelidad e indicadores metacognitivos). Este encuadre robustece la inferencia sobre el impacto y explicita mecanismos plausibles de cambio. La combinación de ciencia abierta —documentación de datasets, scripts y rúbricas— con resguardo ético (anonimización y consentimiento) muestra que es factible sostener estándares editoriales exigentes en investigación educativa aplicada.

No obstante, persisten limitaciones. La asignación por conglomerados deja espacio a confusores no observados; la medida de transferencia, aunque dual (cercana/lejana), no abarca toda la variedad de situaciones del mundo real; y la ventana de seguimiento (8–10 semanas) no captura la trayectoria de decaimiento completa. Estos elementos deben ser abordados en estudios posteriores, idealmente con ensayos aleatorizados por conglomerados, períodos de seguimiento más extensos y protocolos de evaluación ciega por evaluadores externos.

A la luz de los resultados, se derivan implicaciones operativas: (i) incorporar MEAs locales

como tareas fardo en unidades de proporcionalidad, funciones y análisis de datos; (ii) institucionalizar rúbricas de modelación y evaluación formativa con ciclos cortos de retroalimentación; (iii) proveer materiales impresos y datasets resumidos para contextos con conectividad limitada; (iv) mantener prácticas espaciadas y refuerzos que consoliden la ganancia; y (v) formalizar repositorios institucionales para evidencias, scripts y metadatos conforme a principios FAIR.

En síntesis, enseñar Matemática “con acento local” no solo mejora el rendimiento en pruebas, sino que fortalece competencias transferibles para enfrentar problemas no rutinarios dentro y fuera del aula. La originalidad del estudio reside en articular modelación contextual, datos abiertos y andamiaje metacognitivo en el subnivel EGB Superior de una gran ciudad costera, aportando evidencia regional donde las lagunas eran notorias. La contribución es doble: empírica, al estimar efectos robustos y educativamente relevantes; y metodológica, al ofrecer un itinerario reproducible para docentes, directivos y decisores.

Como proyección, el escalamiento sistemático de MEAs en la provincia del Guayas —y su adaptación a otras provincias— puede convertirse en un vector de mejora del razonamiento y la transferencia en la educación básica del país. Para ello, resultan claves la formación situada, la gestión de datos abiertos y el monitoreo de fidelidad e impacto con métricas transparentes. Con estos cimientos, la modelación de problemas del entorno deja de ser un aderezo didáctico y pasa a ser un eje estructurante de la enseñanza de la Matemática en EGB.

6. REFERENCIAS BIBLIOGRÁFICAS

- Ackermans, K., Rusman, E., Nadolski, R., Specht, M., & Brand-Gruwel, S. (2021). Video-enhanced or textual rubrics: Does the Viewbrics’ formative assessment methodology support the mastery of complex (21st century) skills? *Journal of Computer Assisted Learning*, 37(3), 810–824. <https://doi.org/10.1111/jcal.12525>
- Angrist, J. D., & Pischke, J.-S. (2009). *Mostly harmless econometrics: An empiricist’s companion*. Princeton University Press. <https://doi.org/10.1515/9781400829828>
- Barnett, S. M., & Ceci, S. J. (2002). When and where do we apply what we learn? A taxonomy for far transfer. *Psychological Review*, 109(2), 145–170. <https://doi.org/10.1037/0033-295X.109.2.145>
- Black, P., & Wiliam, D. (2018). Classroom assessment and pedagogy. *Assessment in Education: Principles, Policy & Practice*, 25(6), 551–575. <https://doi.org/10.1080/0969594X.2018.1441807>
- English, N., Robertson, P., Gillis, S., & Graham, L. (2022). Rubrics and formative assessment in K-12 education: A scoping review of literature. *International Journal of Educational Research*, 113, 101964. <https://doi.org/10.1016/j.ijer.2022.101964>
- Embretson, S. E., & Reise, S. P. (2000). *Item response theory for psychologists*. Psychology Press. <https://doi.org/10.4324/9781410605269>
- García Cabrera, V. A., Guaman Chimbo, M. E., Rea Minchalo, C. B., & Vega Pérez, J. A. (2025). Impacto de los medios tecnológicos en el aprendizaje de estudiantes de educación básica media en contextos urbanos de Ecuador. *Horizonte Científico International Journal*, 3(2), 1–10. <https://doi.org/10.64747/zvkjx362>
- Greefrath, G., Hopfenbeck, T. N., Lingefjärd, T., & Ludwig, M. (2022). Mathematical modelling and discrete mathematics: Opportunities for modern mathematics teaching.

- ZDM–Mathematics Education, 54(7), 1723–1737. <https://doi.org/10.1007/s11858-022-01339-5>
- Hartmann, L. M., Schukajlow, S., Niss, M., & Jankvist, U. T. (2024). Preservice teachers' metacognitive process variables in modeling-related problem posing. *The Journal of Mathematical Behavior*, 76, 101195. <https://doi.org/10.1016/j.jmathb.2024.101195>
- Hidayat, R., Adnan, M., Abdullah, M. F. N. L., & Safrudiannur. (2022). A systematic literature review of measurement of mathematical modeling in mathematics education context. *Eurasia Journal of Mathematics, Science and Technology Education*, 18(5), em2108. <https://doi.org/10.29333/ejmste/12007>
- Hox, J. J., Moerbeek, M., & van de Schoot, R. (2017). *Multilevel analysis: Techniques and applications* (3rd ed.). Routledge. <https://doi.org/10.4324/9781315650982>
- Imbens, G. W., & Wooldridge, J. M. (2009). Recent developments in the econometrics of program evaluation. *Journal of Economic Literature*, 47(1), 5–86. <https://doi.org/10.1257/jel.47.1.5>
- Lakens, D. (2017). Equivalence tests: A practical primer for t-tests, correlations, and meta-analyses. *Social Psychological and Personality Science*, 8(4), 355–362. <https://doi.org/10.1177/1948550617697177>
- Murillo Cortez, C. A., Muñoz Velastegui, T. del R., Cercado Erazo, M. Á., & Villegas Fajardo, G. A. (2025). Desempeño académico y formación técnica en Guayaquil: Un análisis desde la educación básica hasta el bachillerato técnico. *Horizonte Científico International Journal*, 3(2), 1–13. <https://doi.org/10.64747/tk2xjq24>
- OECD. (2023). *PISA 2022 results (Volume I): The state of learning worldwide*. OECD Publishing. <https://doi.org/10.1787/53f23881-en>
- Pan, S. C., & Rickard, T. C. (2018). Transfer of test-enhanced learning: Meta-analytic review and synthesis. *Current Opinion in Behavioral Sciences*, 20, 20–27. <https://doi.org/10.1016/j.cobeha.2018.02.004>
- Sandefur, J. (2022). Mathematical modelling in school mathematics: Perspectives and next steps. *ZDM–Mathematics Education*, 54(7), 1687–1698. <https://doi.org/10.1007/s11858-022-01399-7>
- Schukajlow, S., Kaiser, G., & Stillman, G. (2023). Modelling and applications in mathematics education—State, trends and issues. *Educational Studies in Mathematics*, 114(2), 219–245. <https://doi.org/10.1007/s10649-023-10265-6>
- Söderström, N. C., & Bjork, R. A. (2015). Learning versus performance: An integrative review. *Perspectives on Psychological Science*, 10(2), 176–199. <https://doi.org/10.1177/1745691615569000>
- van der Linden, W. J. (Ed.). (2017). *Handbook of item response theory: Volume 3—Applications*. Chapman & Hall/CRC. <https://doi.org/10.1201/9781315117430>
- Zöttl, L., Ufer, S., & Reiss, K. (2011). Assessing modelling competencies using a multidimensional IRT approach. In G. Kaiser, W. Blum, R. Borromeo Ferri, & G. Stillman (Eds.), *Trends in the teaching and learning of mathematical modelling* (pp. 427–437). Springer. https://doi.org/10.1007/978-94-007-0910-2_42

Conflicto de Intereses: Los autores declaran que no tienen conflictos de intereses relacionados con este estudio y que todos los procedimientos seguidos cumplen con los estándares éticos establecidos por la revista.

Asimismo, confirman que este trabajo es inédito y no ha sido publicado, ni parcial ni totalmente, en ninguna otra publicación