

Artículo de investigación

<https://doi.org/10.47460/minerva.v7i21.355>

Brecha urbano-rural en el rendimiento matemático: comprensión de problemas y barreras lingüísticas

Jean Carlos Mallco Gayo
<https://orcid.org/0009-0009-1234-2394>
jean.mallco.08@unsch.edu.pe
Universidad Nacional San Cristóbal de Huamanga
Huamanga, Perú

Autor de correspondencia: jean.mallco.08@unsch.edu.pe

Recibido: (12/06/2026), Aceptado: (09/10/2026)

Resumen. Las desigualdades de aprendizaje entre estudiantes de instituciones urbanas y rurales se acentúan cuando la lengua de enseñanza difiere de la lengua materna. Esta investigación analizó las brechas en la resolución de problemas matemáticos y los factores lingüísticos asociados, mediante un estudio transversal explicativo con 395 estudiantes de secundaria de instituciones públicas de Ayacucho, Perú. Se aplicó un análisis factorial confirmatorio y una regresión jerárquica por mínimos cuadrados ordinarios con errores estándar robustos. Los resultados evidenciaron que el contexto rural se asocia con un menor rendimiento matemático, aunque esta desventaja desaparece cuando las dificultades de decodificación lingüística son reducidas y aumenta conforme estas se intensifican. Asimismo, la comprensión de los enunciados matemáticos contribuye a atenuar la desigualdad, en lugar de actuar como mecanismo de transmisión. Se concluye que fortalecer las competencias lingüísticas resulta fundamental para reducir las brechas de aprendizaje entre contextos urbanos y rurales.

Palabras clave: educación rural, bilingüismo quechua-castellano, equidad educativa, evaluación PISA.

Urban–rural gap in mathematics achievement: problem comprehension and linguistic barriers

Abstract. Learning inequalities between students from urban and rural institutions are accentuated when the language of instruction differs from the mother tongue. This research analyzed the gaps in mathematical problem solving and the associated linguistic factors, through an explanatory cross-sectional study with 395 secondary school students from public institutions in Ayacucho, Peru. A confirmatory factor analysis and a hierarchical regression by ordinary least squares with robust standard errors were applied. The results showed that the rural context is associated with lower mathematical performance, although this disadvantage disappears when the difficulties of linguistic decoding are reduced and increases as they intensify. Likewise, the comprehension of mathematical statements contributes to attenuating inequality, instead of acting as a transmission mechanism. It is concluded that strengthening language skills is essential to reduce learning gaps between urban and rural contexts.

Keywords: rural education, Quechua–Spanish bilingualism, educational equity, PISA assessment.

I. INTRODUCCIÓN

La resolución de problemas matemáticos es uno de los indicadores más comparables del aprendizaje escolar y el que revela con mayor nitidez las desigualdades de un sistema educativo. El análisis de tendencias muestra que la distancia entre escuelas urbanas y rurales persiste incluso cuando los sistemas amplían su cobertura [1], y esa regularidad no es local: aparece con magnitudes comparables en el sudeste asiático [2] y en Europa [3], lo que sugiere condiciones estructurales antes que particularidades nacionales.

La brecha adquiere una forma específica cuando la escolarización ocurre en una lengua distinta de la lengua materna. La tarea matemática incorpora entonces una demanda añadida: antes de operar con las cantidades, el estudiante debe reconstruir una situación descrita en una segunda lengua. La evidencia comparada muestra que una menor comprensión lectora en la lengua de instrucción explica buena parte del menor logro matemático de quienes hablan otra lengua en el hogar [4], y una revisión de revisiones sitúa el componente lingüístico entre los predictores más consistentes del desempeño [5]. Cuando a la distancia entre contextos educativos se superpone una distancia lingüística, cabe esperar que la desventaja no solo se suma, sino que cambie de magnitud.

El Perú ofrece un caso informativo de esa superposición: combina una brecha territorial documentada con una diversidad lingüística que la política educativa reconoce mediante la educación intercultural bilingüe, pero que la evaluación estandarizada aún procesa con dificultad [6], [7]. El paisaje lingüístico escolar muestra que el castellano ocupa casi todo el espacio aun donde la lengua originaria es mayoritaria [8], y los estudios sobre juventud quechuahablante describen trayectorias en que la lengua familiar y la escolar se reparten funciones asimétricas [9]. En Ayacucho ambas condiciones coinciden en la misma población.

La pregunta que organiza este estudio no es cuánta brecha hay, sino de qué está hecha: si la desventaja rural se transmite a través de la comprensión del enunciado y si la dificultad para decodificarlo en castellano altera su magnitud. Responderla exige distinguir dos mecanismos que la literatura suele tratar de forma conjunta —uno que transmite el efecto del contexto y otro que lo condiciona— y contrastar cada uno con la prueba que le corresponde [10]. Las tres relaciones examinadas se formulan como hipótesis en el apartado de metodología, junto con la operacionalización de cada variable. Al mostrar que comprensión y decodificación cumplen funciones distintas —una que compensa y otra que amplifica—, el trabajo ofrece un modelo trasladable a cualquier sistema con escolarización en segunda lengua y desplaza el foco de las políticas de dotación hacia las de andamiaje lingüístico: Ayacucho opera aquí como sitio de prueba de una relación general, no como el límite del hallazgo.

II. MARCO TEÓRICO

El rendimiento matemático escolar no es un constructo homogéneo. La resolución de problemas verbales es su faceta más exigente, porque integra el manejo de procedimientos con la interpretación de una situación descrita en lenguaje natural; los marcos de evaluación internacional definen la competencia matemática justamente por esa integración y no por el dominio aislado de algoritmos [11]. Es, por ello, la competencia donde las desigualdades de origen se expresan con mayor claridad.

La brecha educativa urbano-rural en el rendimiento matemático está documentada en evaluaciones nacionales y en estudios comparados [1], [2]. Sus explicaciones habituales combinan la dotación material de los centros, la formación docente, la exposición previa a tareas complejas y las condiciones socioeconómicas del hogar [3]. Lo que esas explicaciones dejan abierto es el mecanismo: por qué camino concreto una condición del entorno se convierte en una diferencia de desempeño individual.

La comprensión de un problema matemático difiere de la lectura general: supone identificar los datos relevantes, reconocer la pregunta y construir una representación de la situación que permita seleccionar una estrategia, y es esa representación la que conecta el texto con la operación [12]. Su calidad predice el éxito en la tarea con relativa independencia de la destreza de cálculo, y la asociación

entre comprensión de enunciados y resolución de problemas se replica en poblaciones escolares diversas [10].

Antes de comprender hay que decodificar. El reconocimiento de palabras, la segmentación de oraciones largas y la interpretación de conectores constituyen un nivel previo cuyo costo pasa inadvertido cuando la lengua del enunciado coincide con la del lector [13]. El análisis de la complejidad lingüística de los problemas escolares y de sus traducciones muestra que ese nivel no es neutro: pequeñas variaciones sintácticas alteran de forma sistemática la dificultad de una misma tarea matemática [14].

En los sistemas andinos, escolarizar en castellano a estudiantes quechuahablantes no equivale a una traducción. Las estructuras de los enunciados, los referentes cotidianos y las convenciones de la prueba provienen de una tradición escolar distinta de la del hogar, de modo que la asimetría no se reduce a vocabulario desconocido, sino que alcanza la familiaridad con el género textual del problema matemático [15]; las trayectorias bilingües descritas para el quechua muestran, además, que esa asimetría se negocia de manera distinta dentro y fuera de la escuela [9].

La teoría de la carga cognitiva explica por qué esa asimetría afecta al rendimiento: la memoria de trabajo es limitada y los recursos destinados a decodificar el enunciado dejan de estar disponibles para el razonamiento matemático [16]. La carga impuesta por la lengua es extrínseca a la tarea y no aporta aprendizaje: solo desplaza capacidad. La evidencia en aulas multilingües confirma que la instrucción en una lengua no dominada opera como carga añadida y que esa carga se traduce en un menor desempeño en problemas verbales [17].

La teoría del aprendizaje situado permite leer el contexto educativo no como una etiqueta geográfica, sino como un entramado de prácticas que define qué cuenta como problema, qué procedimientos son legítimos y con qué recursos se aprende [18]. Bajo esta lectura, la ruralidad no es una carencia en sí misma, sino una configuración de prácticas que puede alinearse mejor o peor con las que la prueba presupone.

El nivel socioeconómico del hogar acompaña casi siempre a la condición territorial y opera como vía alternativa de explicación de la brecha [19]. Por eso, toda estimación del efecto del contexto educativo que aspire a ser interpretable debe ajustarse por él, junto con la lengua materna, para no atribuir al territorio lo que corresponde a la posición social [3].

Si la desventaja rural se explicara por una menor capacidad de comprender los enunciados, la comprensión debería transmitir el efecto del contexto sobre el rendimiento. Esa es la hipótesis de mediación que el estudio somete a prueba, y su valor reside tanto en confirmarse como en refutarse, porque de esa distinción depende qué intervención resulta pertinente: las intervenciones que incorporan instrucción explícita en comprensión del enunciado reducen la brecha en resolución de problemas, lo que hace de la comprensión un candidato plausible a mecanismo [20].

La dificultad para decodificar, en cambio, no compite con el contexto por explicar el rendimiento: opera sobre la relación entre ambos. Si la carga lingüística consume los recursos que el razonamiento requiere, la desventaja rural debería ser mayor entre quienes enfrentan una barrera alta y menor —o nula— entre quienes no la enfrentan. Los análisis de grandes evaluaciones que distinguen la condición de minoría lingüística apuntan en esa dirección [21]. Es una hipótesis sobre la forma de la relación, no sobre su existencia, y exige un término de interacción para contrastarse.

III. METODOLOGÍA

A. *Diseño e hipótesis*

La investigación adoptó un enfoque cuantitativo, con diseño transversal y alcance explicativo, considerando al estudiante como unidad de análisis. La estrategia estadística comprendió un análisis factorial confirmatorio para evaluar el modelo de medición, seguido de una regresión jerárquica por mínimos cuadrados ordinarios, utilizando las puntuaciones de los constructos. Este procedimiento permitió exa-

minar las relaciones entre el contexto educativo, las barreras lingüísticas y el rendimiento en la resolución de problemas matemáticos, así como evaluar el efecto moderador de las dificultades de decodificación lingüística. A partir de los fundamentos teóricos se formularon dos hipótesis de investigación:

H1. El contexto educativo rural se asocia negativamente con el rendimiento en la resolución de problemas matemáticos, de manera que los estudiantes de instituciones rurales presentan un desempeño inferior al de sus pares de instituciones urbanas.

H2. La barrera de decodificación lingüística modera la relación entre el contexto educativo y el rendimiento en la resolución de problemas matemáticos, de modo que la desventaja asociada con la ruralidad se intensifica a medida que aumentan las dificultades para comprender e interpretar los enunciados matemáticos.

B. Participantes y procedimiento de muestreo

Durante el año escolar 2026, participaron 395 estudiantes de cuarto grado de secundaria de instituciones educativas públicas de la región de Ayacucho, Perú, de los cuales 210 pertenecían a instituciones rurales y 185 a instituciones urbanas. La determinación del tamaño muestral consideró los requerimientos estadísticos del modelo de regresión, particularmente la identificación del efecto de interacción entre el contexto educativo y las barreras de decodificación lingüística. Para ello, se establecieron como criterios una potencia estadística de 0,80, un nivel de significación de 0,05 y un tamaño del efecto mínimo de interés de $f^2 = 0,02$. Asimismo, se consideraron las exigencias del análisis factorial confirmatorio, con el propósito de disponer de un número suficiente de observaciones para evaluar el modelo de medición y contrastar las hipótesis propuestas.

La muestra procedió de 61 instituciones educativas públicas, distribuidas entre 31 urbanas y 30 rurales, con un promedio aproximado de 6,5 estudiantes por institución. Se empleó un muestreo no probabilístico e intencional, considerando la accesibilidad de los participantes y la inclusión de ambos contextos educativos. Los criterios de inclusión comprendieron la matrícula vigente en el grado correspondiente, la asistencia regular y la autorización mediante consentimiento informado del representante legal. Se excluyeron los registros con información incompleta en las variables necesarias para el análisis. Debido al procedimiento de selección, los resultados se circunscriben a los estudiantes participantes y no permiten establecer inferencias estadísticamente representativas de toda la población escolar de Ayacucho.

Considerando que los estudiantes pertenecían a diferentes instituciones educativas, se reconoció la posibilidad de dependencia estadística entre las observaciones de un mismo centro. La distribución de la muestra entre numerosas instituciones, con una participación relativamente reducida por establecimiento, permitió disminuir la concentración de observaciones, aunque no eliminó el posible efecto de agrupamiento. Como aproximación, se examinaron escenarios hipotéticos de correlación intraclass de 0,05, 0,10 y 0,15, que produjeron efectos de diseño estimados de 1,27, 1,55 y 1,82, respectivamente, y tamaños muestrales efectivos aproximados de 310, 255 y 217 observaciones. Estos cálculos se utilizaron como referencia para valorar la sensibilidad del tamaño muestral ante distintos niveles de dependencia. El análisis se desarrolló a nivel individual mediante errores estándar robustos frente a la heterocedasticidad. No obstante, dado que este procedimiento no corrige por sí mismo la correlación entre estudiantes de una misma institución, la ausencia de un ajuste específico por agrupamiento se reconoce como una limitación metodológica del estudio.

C. Variables e instrumentos

El contexto educativo urbano-rural es una variable observada dicotómica, definida por la clasificación oficial del centro educativo y codificada Urbano = 0 y Rural = 1; esa codificación gobierna el signo de todos los coeficientes en que interviene. El rendimiento se midió con reactivos liberados del cuadernillo PISA 2022, en adaptación autorizada, correspondientes a los cuatro dominios de contenido evaluados [11]; la ejecución en cada dominio se calificó como acierto o error y el puntaje bruto se equiparó a los cinco niveles de logro ordenados del currículo nacional, de modo que el constructo se expresa en una

escala ordinal de cinco categorías y no en una escala de autorreporte. La comprensión de problemas y la barrera de decodificación se midieron como constructos latentes reflectivos de cuatro reactivos cada uno, con escala Likert de cinco puntos, apoyados respectivamente en el modelo de comprensión de problemas verbales [12] y en una escala de dificultad de decodificación de elaboración propia. Como covariables de control se incluyeron el nivel socioeconómico del hogar, índice compuesto de cuatro reactivos de categorías ordenadas, y la lengua materna quechua, variable dicotómica autorreportada.

D. Estrategia analítica

El análisis se desarrolla en un único marco, la regresión sobre puntuaciones de constructos previamente validados, y cada procedimiento cumple una función delimitada. El análisis factorial confirmatorio verifica que los reactivos midan los tres constructos declarados antes de que entren en el modelo explicativo, y de él proceden los índices de ajuste, las cargas, las fiabilidades y la varianza media extraída que se reportan; ninguno de esos índices describe al modelo de regresión. El contraste con un modelo unifactorial y el análisis de invarianza entre contextos comprueban, respectivamente, que los constructos son empíricamente distinguibles y que su medición es comparable entre estudiantes urbanos y rurales, condición sin la cual comparar coeficientes carecería de sentido.

La regresión jerárquica estima el efecto del contexto y su interacción con la barrera en cuatro pasos: M1 introduce las covariables de control; M2 añade el contexto educativo; M3 añade el mediador y el moderador; y M4 añade el término de interacción. Los errores estándar son robustos a la heterocedasticidad. El efecto indirecto se estima con intervalos de confianza por bootstrap de 5000 remuestras, y los efectos condicionales, mediante pendientes simples evaluadas en la media del moderador y a ± 1 desviación estándar. Por último, las pruebas de estrés por jackknife, submuestreo y remuestreo Monte Carlo evalúan la estabilidad del coeficiente focal.

IV. RESULTADOS

Los resultados se presentan en cinco capas: modelo de medición, modelo estructural, descomposición de efectos, efectos condicionales y robustez.

A. Modelo de medición

Los tres constructos alcanzan cargas factoriales por encima del criterio en todos sus reactivos y fiabilidades holgadamente suficientes, con varianzas medias extraídas superiores al 50 % (Tabla 1). El modelo de tres factores ajusta bien la matriz de covarianzas observada, mientras que la alternativa unifactorial resulta inadmisibles, y la invarianza de medición entre contextos se sostiene en los pasos evaluados, de modo que las diferencias que el modelo estructural atribuye al contexto no son artefactos de una medición que funcione distinto en cada grupo: el modelo de tres factores alcanza $CFI = 0,993$, $TLI = 0,991$, $RMSEA = 0,026$ y $SRMR = 0,053$, frente a $CFI = 0,359$ y $RMSEA = 0,237$ de la alternativa unifactorial, y el contraste secuencial de invarianza entre contextos se sostiene con $|\Delta CFI| \leq 0,001$ y $|\Delta RMSEA| \leq 0,006$, por debajo de los criterios convencionales de 0,010 y 0,015. La validez discriminante se confirma por la razón heterorrasgo-monorrasgo, cuyo máximo (0,262) queda muy por debajo del umbral estricto de 0,85. La Tabla 1 presenta la carga de cada uno de los doce reactivos frente al criterio de aceptación.

Tabla 1. Modelo de medición: propiedades psicométricas por reactivo

Constructo	Reactivo	λ	R^2 (SMC)	α	ω	AVE
Rendimiento	Problema de cambio y relaciones	0,709	0,503	0,834	0,835	0,560
	Problema de espacio y forma	0,782	0,612			
	Problema de cantidad	0,818	0,669			
	Problema de incertidumbre y datos	0,677	0,458			
Barrera de decodificación	Me cuesta entender el significado de las palabras en los problemas	0,797	0,635	0,818	0,819	0,532
	Cuando leo un problema, no sé qué operación debo realizar	0,696	0,484			
	Los enunciados largos me confunden y no sé por dónde empezar	0,708	0,501			
	Necesito que me expliquen el problema en quechua para entenderlo	0,712	0,507			
Comprensión de problemas	Puedo identificar los datos importantes de un problema	0,788	0,621	0,861	0,863	0,613
	Reconozco la pregunta que debo responder en cada problema	0,839	0,704			
	Puedo representar el problema con mis propias palabras	0,824	0,679			
	Distingo la información relevante de la irrelevante en el enunciado	0,670	0,449			

B. Modelo estructural

El patrón de asociaciones bivariadas (Tabla 2) anticipa la estructura del modelo: el contexto se asocia negativamente con el rendimiento y positivamente con la comprensión y con la barrera, mientras que la lengua materna quechua no se asocia con el desempeño. La incorporación del contexto produce el mayor salto de varianza explicada, y su coeficiente se mantiene negativo y significativo en los tres bloques en que interviene: los estudiantes de instituciones rurales rinden por debajo de sus pares urbanos una vez descontados el nivel socioeconómico y la lengua materna (Tabla 3). El nivel socioeconómico conserva un efecto positivo pequeño y significativo, mientras que la lengua materna quechua, por sí sola, no discrimina el rendimiento: lo que pesa no es hablar quechua, sino la dificultad efectiva para decodificar el enunciado, que sí lo hace. El término de interacción, incorporado en M4, es negativo y significativo y eleva la varianza explicada, con factores de inflación de la varianza por debajo del umbral convencional.

Tabla 2. Estadísticos descriptivos y matriz de correlaciones

N.º	Variable	M	DE	1	2	3	4	5
1	Contexto educativo (Rural = 1)	0,532	0,500	—				
2	Rendimiento en resolución de problemas	3,001	0,663	−0,297***	—			
3	Comprensión de problemas	3,000	0,667	0,300***	0,063	—		
4	Barrera de decodificación	3,000	0,667	0,294***	−0,344***	0,043	—	
5	Nivel socioeconómico	2,996	0,651	0,031	0,119*	0,033	−0,001	—
6	Lengua materna quechua	0,476	0,500	−0,020	0,015	−0,076	−0,092	0,003

La Tabla 3 presenta los resultados de la regresión jerárquica por mínimos cuadrados ordinarios con errores estándar robustos, desarrollada mediante cuatro modelos sucesivos para examinar los factores asociados con el rendimiento en la resolución de problemas matemáticos. Se observa un incremento progresivo de la varianza explicada, desde el 1,4 % en el modelo inicial hasta el 27,9 % en el modelo final, lo que evidencia una mayor capacidad explicativa al incorporar las variables educativas y lingüísticas. El contexto rural mantiene una asociación negativa y estadísticamente significativa con el rendimiento, mientras que la comprensión de problemas presenta una relación positiva. Asimismo, el término de interacción entre el contexto educativo y la barrera de decodificación muestra un coeficiente negativo

significativo ($\beta = -0,295$; $p < 0,001$), lo que indica que las dificultades lingüísticas intensifican la desventaja asociada con la ruralidad. Estos resultados respaldan las hipótesis planteadas y destacan la importancia de considerar la interacción entre el contexto educativo y las barreras lingüísticas para comprender las desigualdades en el aprendizaje matemático.

Tabla 3. Regresión jerárquica por mínimos cuadrados ordinarios con errores robustos

Predictor	M1	M2	M3	M4	VIF (M4)
Nivel socioeconómico	0,021* [0,005; 0,038]	0,023** [0,007; 0,039]	0,022** [0,007; 0,037]	0,027*** [0,013; 0,041]	1,01
Lengua materna quechua	0,014 [−0,082; 0,111]	0,009 [−0,083; 0,101]	−0,004 [−0,092; 0,085]	−0,007 [−0,092; 0,077]	1,01
Contexto educativo (Rural = 1)	—	−0,293*** [−0,384; −0,202]	−0,259*** [−0,355; −0,163]	−0,272*** [−0,360; −0,184]	1,10
Barrera de decodificación (est.)	—	—	−0,133*** [−0,181; −0,085]	0,032 [−0,031; 0,096]	2,52
Comprensión de problemas	—	—	0,190** [0,059; 0,320]	0,171** [0,047; 0,296]	—
Contexto × Barrera	—	—	—	−0,295*** [−0,382; −0,208]	2,39
R^2 del modelo	0,014	0,105	0,197	0,279	—

C. Descomposición de efectos

El análisis del efecto indirecto ofrece el resultado que más altera la lectura habitual de esta brecha (Tabla 4). La ruta del contexto a la comprensión es positiva, la de la comprensión al rendimiento también, y el efecto directo del contexto sobre el rendimiento es negativo: el producto de las dos primeras, que constituye el efecto indirecto, resulta positivo y de signo contrario al directo. La configuración no corresponde a una mediación parcial, en la que el mediador transmitiría parte de la desventaja, sino a una mediación competitiva o de supresión, en la que el mediador opera en dirección opuesta y atenúa el efecto total. La desventaja rural no viaja, por tanto, a través de la comprensión de los enunciados: la comprensión actúa en esta población como vía compensatoria de magnitud modesta pero consistente.

La lectura es coherente con la comparación de dos especificaciones anidadas: el coeficiente del contexto pasa de $-0,2304$, cuando el mediador se excluye del modelo M4, a $-0,272$ cuando se incluye. Que la magnitud aumente al introducir el mediador es el comportamiento propio de una supresión, y no el de una mediación parcial, en la que el efecto directo se reduciría. El coeficiente c' de la Tabla 4 pertenece al modelo de mediación, que no incorpora el término de interacción, y por eso no es comparable cifra a cifra con los de la regresión jerárquica. La Tabla 4 representa cada ruta de esta descomposición con su intervalo de confianza.

Tabla 4. Descomposición de efectos: rutas directa, indirecta y total

Ruta	Coefficiente	IC 95 %	Significativa
<i>a</i> : contexto → comprensión	0,232	[0,162; 0,302]	Sí
<i>b</i> : comprensión → rendimiento	0,211	[0,080; 0,344]	Sí
<i>c'</i> : efecto directo del contexto sobre el rendimiento	−0,338	[−0,432; −0,250]	Sí
<i>a</i> × <i>b</i> : efecto indirecto	0,049	[0,018; 0,086]	Sí
<i>c</i> : efecto total (<i>c'</i> + <i>a</i> × <i>b</i>)	−0,289	[−0,381; −0,199]	Sí

D. Efectos condicionales

La desventaja del contexto rural no es una cantidad fija: varía con el nivel del moderador según la combinación de los coeficientes de M4, esto es, el coeficiente del contexto más el de la interacción por el nivel estandarizado de la barrera. En el nivel bajo la diferencia entre contextos no difiere significativamente de cero; en el nivel medio se establece con claridad; y en el alto duplica con creces la magnitud del nivel medio: $\beta = 0,023$ (EE = 0,063; $p = 0,711$) en el nivel bajo, $\beta = -0,272$ (EE = 0,045; $p < 0,001$) en la media y $\beta = -0,567$ (EE = 0,063; $p < 0,001$) en el alto, los tres consistentes por construcción con el coeficiente de interacción de $-0,295$. La brecha se amplifica, por tanto, a medida que aumenta la dificultad para decodificar el enunciado, y no se atenúa. El hallazgo delimita el alcance de H1: no existe una desventaja rural uniforme, sino una que se activa cuando la demanda lingüística de la tarea es alta. La Figura 1 muestra el efecto condicional a lo largo del moderador y el nivel en que cruza el cero.

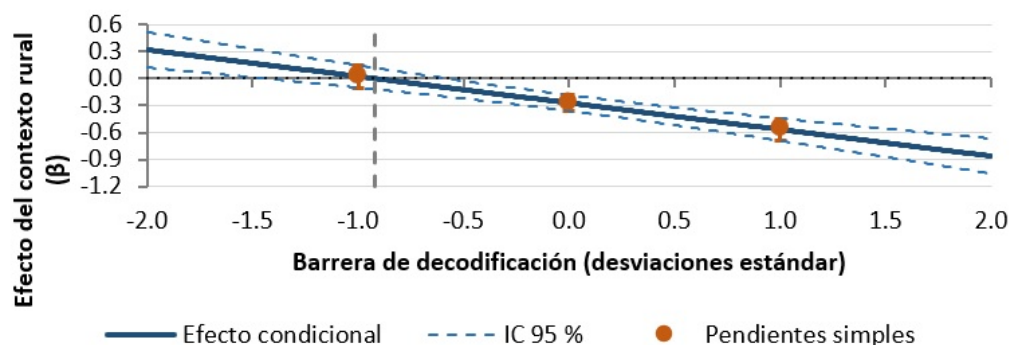


Fig. 1. Efecto condicional del contexto rural sobre el rendimiento matemático según el nivel de barrera de decodificación lingüística

E. Robustez

El coeficiente del contexto conserva signo y significación en la práctica totalidad de los reajustes de las cuatro pruebas de perturbación muestral, con una dispersión reducida entre remuestras: jackknife sobre 395 reajustes, β medio = $-0,2304$ (DE = 0,0022); submuestreo al 80 % y al 60 % sobre 100 réplicas, $-0,2308$ y $-0,2328$; y bootstrap de Monte Carlo con 1000 réplicas, $-0,2284$. El signo se conserva en el 100 % de los reajustes de las cuatro pruebas. La estabilidad no debe leerse como validación externa del modelo, sino como evidencia de que el resultado no depende de observaciones particulares de la muestra analizada. La Figura 2 resume la estabilidad del coeficiente focal en las cuatro pruebas. El coeficiente sometido a estas pruebas es el del contexto en la especificación de M4 sin el mediador (β base = $-0,2304$), que difiere del contexto en el M4 completo ($-0,272$) porque la entrada del mediador aumenta la magnitud del efecto directo, comportamiento propio de una supresión.

En síntesis, H1 y H2 se sostienen —el coeficiente del contexto es negativo y significativo en todos los bloques en que interviene, y el término de interacción lo es también—, mientras que H3 no se sostiene en la dirección hipotetizada: el efecto indirecto es significativo, pero de signo opuesto al efecto directo y al postulado, de modo que la comprensión no transmite la desventaja rural, sino que la compensa

parcialmente. La relación se reporta como mediación competitiva; presentarla como mediación parcial contradiría los coeficientes estimados.

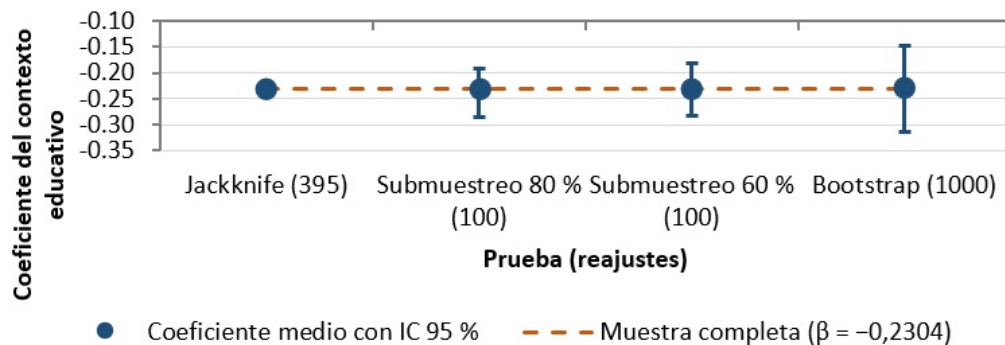


Fig. 2. Estabilidad del coeficiente del contexto educativo ante cuatro pruebas de perturbación de la muestra (N = 395)

F. Discusión

La brecha rural en el rendimiento matemático persiste una vez incorporados al modelo el mediador, el moderador y las covariables de control, en línea con la literatura sobre desigualdades territoriales en sistemas de renta media [1], [2], [19]. El hallazgo central, sin embargo, es que esa desventaja no es uniforme: se anula cuando la barrera de decodificación es baja y se amplifica cuando es alta. El contexto rural no opera, por tanto, como una condición homogénea, sino como un espacio donde la distancia entre lengua materna y lengua de escolarización actúa de filtro diferencial: quienes dominan el castellano escrito sortean el obstáculo; quienes no, acumulan una desventaja que el sistema escolar no compensa. La dirección coincide con la de los análisis internacionales que identifican la condición de minoría lingüística como moderador del logro.

El resultado más contraintuitivo es el de la mediación. La comprensión del enunciado no transmite la desventaja rural: opera en dirección contraria y la atenúa parcialmente. Este patrón de supresión indica que la comprensión aporta un beneficio en ambos contextos, pero no es la vía por la que el grupo rural queda rezagado. El mecanismo decisivo no es cuánto se comprende, sino qué fracción del estudiantado rural llega a la instancia donde la comprensión podría nivelar diferencias: la barrera lingüística actúa antes, en la decodificación. La lectura es compatible con la evidencia experimental según la cual instruir explícitamente en comprensión del enunciado reduce la brecha, pero solo una vez asegurado el acceso al texto [20].

El hallazgo conecta dos marcos que rara vez dialogan: la teoría de la carga cognitiva, que predice que la decodificación consume recursos destinados al procesamiento matemático [16], [17], y la socio-lingüística educativa, que documenta cómo la lengua de escolarización reproduce jerarquías territoriales [15]. Su cruce precisa que la intervención no debe centrarse solo en enseñar mejor matemáticas, sino en asegurar que el enunciado sea accesible antes de exigir su resolución; la magnitud del efecto de la competencia lingüística sobre el desempeño matemático hallada en otros sistemas respalda esa prioridad [5]. En términos aplicados, los programas de fortalecimiento del castellano académico con énfasis en vocabulario matemático deberían implementarse antes de cuarto de secundaria, que es el grado en que este estudio detecta la brecha ya consolidada; los materiales bilingües que presentan el problema primero en la lengua materna actúan sobre la decodificación sin reducir la exigencia cognitiva de la tarea, y las pedagogías de apoyo lingüístico diseñadas para aulas multilingües ofrecen procedimientos ya probados.

CONCLUSIONES

El aporte teórico de este estudio es mostrar que la brecha urbano-rural en matemáticas no es una desventaja uniforme del contexto, sino un efecto condicionado por la barrera de decodificación lingüística, que determina quién puede beneficiarse de la comprensión del enunciado como mecanismo igualador. La distinción importa porque separa dos procesos que la literatura suele tratar como uno solo: el que transmite la desventaja y el que la modula. Aquí, el primero no opera en la dirección esperada y el segundo resulta decisivo.

De ello se siguen tres implicaciones prácticas. La intervención lingüística debe anteceder a la matemática y situarse en los grados previos a aquel en que la brecha se mide, con seguimiento periódico por la instancia local de gestión educativa. Los materiales destinados a escuelas rurales deberían incorporar glosarios bilingües quechua-castellano en los propios enunciados, reduciendo la carga de decodificación sin eliminar la exigencia cognitiva de la tarea. Y las evaluaciones regionales deberían medir por separado decodificación y comprensión, porque este estudio muestra que cumplen funciones distintas y que agregarlas oculta el mecanismo que produce la brecha.

Estas conclusiones describen a la población efectivamente analizada y no reclaman representatividad estadística de la región. La investigación futura debería evaluar si la intervención lingüística temprana reduce la brecha medida al final de la secundaria, comparando cohortes expuestas y no expuestas; examinar si el patrón se replica en otras competencias; y explorar si la barrera de decodificación interactúa también con condiciones de la enseñanza. La educación intercultural bilingüe ofrece el marco institucional para ensayar esas respuestas.

REFERENCIAS

- [1] E. Mohammadpour and H. Yon, "Mathematics achievement at rural and urban secondary schools: a trends analysis," *Mathematics Education Research Journal*, vol. 37, no. 4, pp. 709–742, 2024.
- [2] R. Minami and K. Ogawa, "Mind the gap: uncovering student's learning achievement disparities in cambodian urban and rural secondary schools," *International Journal of Educational Development*, vol. 118, p. 103420, 2025.
- [3] D. Bianchi and L. Cabrera, "El rendimiento académico del alumnado de 4.º de primaria de origen rural/urbano en España: el papel de la brecha digital y la desigualdad socioeconómica," *Bordón. Revista de Pedagogía*, vol. 78, no. 1, pp. 71–89, 2026.
- [4] M. Greisen, C. Georges, C. Hornung, P. Sonnleitner, and C. Schiltz, "Learning mathematics with shackles: how lower reading comprehension in the language of mathematics instruction accounts for lower mathematics achievement in speakers of different home languages," *Acta Psychologica*, vol. 221, p. 103456, 2021.
- [5] M. Geçici, F. Geçici, and M. Kaya, "Language and mathematics performance: a systematic umbrella review and meta-meta-analysis," *Journal of Intelligence*, vol. 14, no. 8, p. 152, 2026.
- [6] N. Fernandez, "Educación intercultural bilingüe en el Perú: desafíos de una deuda histórica con los pueblos originarios," *Amazonía Peruana*, no. 39, pp. 127–158, 2026.
- [7] J. Tarrillo, L. Mamani, E. Tocto, H. Iftikhar, R. Gonzales, and J. Jesús, "Exploring patterns in intercultural bilingual education in Peru using machine learning," *Frontiers in Education*, vol. 11, 2026.
- [8] K. Sullón, "El paisaje lingüístico en las escuelas de educación intercultural bilingüe del Perú," *Revista Iberoamericana de Educación*, vol. 96, no. 1, pp. 83–95, 2024.
- [9] F. Kvietok, "Migrant bilingual youth, family, and school language policy: ethnographic insights for urban quechua education," *International Journal of the Sociology of Language*, vol. 2023, no. 280, pp. 143–166, 2023.
- [10] F. Pascua and M. Azares, "Reading comprehension and mathematical problem-solving skills of grade 10 students: a correlational study," *Divine Word International Journal of Management and Humanities*, vol. 5, no. 2, pp. 3891–3909, 2026.

- [11] OECD, *PISA 2022 Assessment and Analytical Framework*. Paris, France: OECD Publishing, 2023.
- [12] W. Kintsch and J. Greeno, "Understanding and solving word arithmetic problems," *Psychological Review*, vol. 92, no. 1, pp. 109–129, 1985.
- [13] W. Kintsch and T. van Dijk, "Toward a model of text comprehension and production," *Psychological Review*, vol. 85, no. 5, pp. 363–394, 1978.
- [14] L. Domke, M. Cerrato, E. Sanders, and M. Vo, "Linguistic complexity of elementary mathematics word problems and their translations," *Language and Education*, vol. 39, no. 1, pp. 37–53, 2024.
- [15] J. Hinojosa, Y. Perez, C. Tintaya, V. Flores, and F. Chalco, "Bilingual intercultural education and the construction of quechua-aymara citizenship among university students in the altiplano region," *Universidad Ciencia y Tecnología*, vol. 28, no. 123, pp. 122–130, 2024.
- [16] J. Sweller, "Cognitive load during problem solving: effects on learning," *Cognitive Science*, vol. 12, no. 2, pp. 257–285, 1988.
- [17] M. Keevy and M. Tharapos, "When language becomes load: English medium instruction and cognitive demands in multilingual accounting classrooms," *Accounting Education*, pp. 1–28, 2026.
- [18] J. Lave and E. Wenger, *Situated Learning: Legitimate Peripheral Participation*. Cambridge, UK: Cambridge University Press, 1991.
- [19] L. Remache and F. Auqui, "Influencia del contexto socioeconómico en el rendimiento académico: rural vs. urbano en alausí-chunchi 2021-2023," *Arandu UTIC*, vol. 12, no. 1, pp. 87–106, 2025.
- [20] L. Fuchs, P. Seethaler, S. Sterba, C. Craddock, D. Fuchs, D. Compton, D. Geary, and P. Changas, "Closing the word-problem achievement gap in first grade: schema-based word-problem intervention with embedded language comprehension instruction," *Journal of Educational Psychology*, vol. 113, no. 1, pp. 86–103, 2021.
- [21] Y. Wu and M. Davis, "Rethinking achievement gaps: language minority status and audience-specific heritage language use in pisa 2018," *Journal of Multilingual and Multicultural Development*, pp. 1–17, 2026.