

DOI:

Type of article: Original, Review, Case Report, Short communication, Academic essay.

Randomized controlled trials (RCTs) in the educational sciences in Latin America

Ensayos controlados aleatorizados (rcts) en las ciencias de la educación en latinoamérica

Jose Humberto Puente, Servicios Múltiples Profesionales, Maturin, Venezuela
<https://orcid.org/0009-0006-0100-9404>, josepuente67@gmail.com

Corresponding author: josepuente67@gmail.com

Citar como: Puente, J. H. (2024). Ensayos controlados aleatorizados (rcts) en las ciencias de la educación en latinoamérica. *International Journal of Innovation and Pedagogy*, 1(1), 20-38. <https://editorialdidaxis.com/ijip/index.php/home/article/view/13>

Recibido: 25-09-2024

Aceptado: 13-10-2024

Publicado: 30-12-2024

Editor: Mgtr. Daniel Roman-Acosta PhD. (c)

ABSTRACT

Randomized controlled trials (RCTs) have emerged as a methodology of increasing relevance in the educational sciences, particularly following the international recognition bestowed with the 2019 Nobel Prize in Economics upon Abhijit Banerjee, Esther Duflo, and Michael Kremer. This article presents a comprehensive analysis of the paradigmatic positioning of RCTs within educational research in Latin America, examining their epistemological foundations, their practical application, and the critiques formulated from various theoretical perspectives. Through a systematic literature review, the tension between the positivist-postpositivist paradigm, which provides the philosophical underpinnings for RCTs, and the socio-critical paradigm, predominant in the Latin American research tradition, is analyzed. The results reveal that, while RCTs offer rigorous evidence for specific and limited interventions, they present significant limitations in capturing the complexity of educational processes in contexts of structural inequality. It is concluded that methodological integration, rather than exclusive adherence to a paradigm, constitutes the most promising path to advance towards educational research that is simultaneously rigorous, contextually relevant, and socially committed.

Keywords: randomized controlled trials, educational research paradigms, Latin America, positivism, socio-critical paradigm, impact evaluation, educational policies.

RESUMEN

Los ensayos controlados aleatorizados (RCTs, por sus siglas en inglés) han emergido como una metodología de creciente relevancia en las ciencias de la educación, particularmente tras el reconocimiento internacional otorgado con el Premio Nobel de Economía 2019, a Abhijit Banerjee, Esther Duflo y Michael Kremer. Este artículo presenta un análisis exhaustivo del

posicionamiento paradigmático de los RCTs dentro de la investigación educativa en Latinoamérica, examinando sus fundamentos epistemológicos, su aplicación práctica y las críticas que desde diversos enfoques teóricos se han formulado. Mediante una revisión documental sistemática, se analiza la tensión entre el paradigma positivista-postpositivista, que proporciona el sustento filosófico a los RCTs, y el paradigma socio-crítico, predominante en la tradición investigativa latinoamericana. Los resultados revelan que, si bien los RCTs ofrecen evidencia rigurosa para intervenciones específicas y acotadas, presentan limitaciones significativas para capturar la complejidad de los procesos educativos en contextos de desigualdad estructural. Se concluye que la integración metodológica, más que la adhesión exclusiva a un paradigma constituye la vía más prometedora para avanzar hacia una investigación educativa que sea simultáneamente rigurosa, contextualmente relevante y socialmente comprometida.

Palabras clave: ensayos controlados aleatorizados, paradigmas de investigación educativa, Latinoamérica, positivismo, paradigma socio-crítico, evaluación de impacto, políticas educativas.

Copyright: © 2024 Autor Jose Humberto Puente; Este es un artículo de acceso abierto distribuido bajo los términos de la licencia de uso y distribución Creative Commons Reconocimiento 4.0 Internacional (CC BY 4.0) 

INTRODUCCIÓN

La concesión del Premio Nobel de Economía a investigadores que promueven el enfoque experimental marcó un punto de inflexión decisivo en la forma de abordar los desafíos del desarrollo global. Brookings Institution (2019), señala que este reconocimiento alteró la cultura de la disciplina, priorizando el rigor empírico sobre la especulación teórica. Duflo, Banerjee y Kremer (2020), describen esta evolución como el auge de una economía del desarrollo moderna donde la evidencia impulsa las políticas. De manera similar, el Banco Mundial (2019), posicionó este galardón como una validación de los ensayos controlados aleatorizados como herramientas esenciales para combatir la pobreza. Sin embargo, el resumen oficial del Premio Nobel (2019), sugiere que, si bien este método avanza en la comprensión del desarrollo, también establece una jerarquía metodológica específica que privilegia los datos experimentales, una tendencia que posteriormente permeó otras ciencias sociales, incluida la educación.

La institucionalización de esta metodología se ejemplifica en el Abdul Latif Jameel Poverty Action Lab, el cual aboga por los ensayos controlados aleatorizados como mecanismo principal de evaluación de políticas (J-PAL, 2024). Los defensores argumentan que estos diseños ofrecen protección contra estimaciones sesgadas de los efectos del tratamiento, imitando el rigor de los entornos clínicos (Cochrane, 2011). A pesar de esto, la aplicación de métodos de "estándar de oro" en educación enfrenta un escrutinio significativo. Connolly, Keenan y Urbanska (2018), documentan la proliferación de estos ensayos, pero Lortie-Forgues e Inglis (2019), advierten que los ensayos a gran escala a menudo arrojan resultados poco informativos debido a la variabilidad contextual. Además, Parra y Edwards (2024), cuestionan el consenso en torno a estos ensayos, destacando los peligros de ignorar el contexto en la educación basada en evidencias.

La imposición de modelos experimentales genera fricción con las tradiciones socio-críticas establecidas, especialmente en regiones donde la educación se concibe como una herramienta de liberación. Por ejemplo, Freire (1970), estableció un marco donde el diálogo y la conciencia crítica son fundamentales, contrastando agudamente con la observación distanciada que requieren los diseños experimentales. Como explica UNAM (2004), los paradigmas de investigación dictan no solo los métodos sino la visión ontológica de la realidad; así, un enfoque positivista choca con la comprensión dialéctica de la educación. Hernández, Fernández y Baptista (2018), distinguen entre rutas cuantitativas y cualitativas, señalando que la elección implica diferentes concepciones del conocimiento. Kuhn (1962), argumentaría que esto representa un choque entre paradigmas inconmensurables, donde la estructura de las revoluciones científicas dificulta la coexistencia entre la lógica objetivista de Popper (1959) y la intención transformadora de la educación popular.

En consecuencia, analizar la alineación paradigmática de estos ensayos revela una base fundamentalmente post-positivista que busca leyes de causalidad generalizables. Gorard, See y Siddiqui (2017), analizan las pruebas de la educación basada en evidencias, enfatizando la estricta metodología que caracteriza este enfoque. Sin embargo, Styles y Torgerson (2018), reconocen debates continuos sobre si este rigor captura la complejidad de los procesos educativos. Kaplan et al. (2020), abogan por redefinir la práctica basada en evidencias para incluir el papel del contexto, ya que los hallazgos a menudo fallan al traducirse entre entornos distintos. Shalem y De Clercq (2022), cuestionan aún más la contribución de los ensayos a la teoría educativa, sugiriendo que pueden explicar "qué funciona" sin abordar adecuadamente el "por qué". Por ello, como sugieren Smith y Flanagan (2024), repensar estos ensayos es necesario para equilibrar el rigor empírico con los matices contextuales.

Marco Teórico: Paradigmas de Investigación Educativa

El Paradigma Positivista y su Evolución hacia el Postpositivismo

El paradigma positivista constituye el fundamento epistemológico de la metodología experimental, asumiendo una realidad objetiva medible a través de la observación empírica. Hernández y Mendoza (2023), identifican esta postura con la ruta cuantitativa, donde la neutralidad valorativa y el control riguroso de variables son primordiales para establecer leyes generales. En este contexto, los ensayos controlados aleatorizados (RCTs) surgieron como herramientas diseñadas para minimizar el sesgo y determinar causalidades, tal como señala Cochrane (2011), al describirlos como una protección contra estimaciones distorsionadas. La influencia de este enfoque se consolidó con figuras como Popper (1959), quien enfatizó la falsificación como criterio de demarcación científica. Sin embargo, la

evolución hacia el postpositivismo, descrita por Kuhn (1962), reconoció que el conocimiento científico es provisional y sujeto a cambios paradigmáticos. A pesar de esta evolución, defensores como Duflo et al. (2020) y el Premio Nobel (2019), sostienen que la experimentalidad rigurosa sigue siendo la vía óptima para generar evidencia válida, posicionando a los RCTs como un estándar metodológico para comprender fenómenos complejos mediante la replicación y el análisis estadístico (Styles & Torgerson, 2018).

El Paradigma Interpretativo y la Comprensión Hermenéutica

En contraposición al enfoque empirista, el paradigma interpretativo o constructivista postula que la realidad social es subjetiva y se construye activamente a través de significados compartidos. Hernández, Fernández y Baptista (2018), señalan que esta perspectiva prioriza la comprensión profunda de los fenómenos en sus contextos naturales por encima de la predicción y el control estadístico. Desde esta óptica, la investigación educativa debe centrarse en la experiencia del sujeto, aspectos que, según Rowe et al. (2016), son a menudo limitados o ignorados por las restricciones ontológicas de los RCTs. Kaplan et al. (2020), argumentan que ignorar el rol del contexto en las evaluaciones experimentales puede llevar a conclusiones erróneas sobre la práctica basada en evidencia. Asimismo, Shalem y De Clercq (2022), cuestionan si los diseños experimentales logran contribuir sustancialmente a la teoría pedagógica o si solo describen efectos superficiales. Román (2024), enfatiza que la terminología y metodología cualitativa son esenciales para capturar la riqueza educativa que los métodos estrictamente cuantitativos tienden a simplificar.

El Paradigma Socio-Crítico en la Tradición Latinoamericana

El paradigma socio-crítico, con fuerte arraigo en América Latina, desafía la supuesta neutralidad de la ciencia positivista y busca la transformación de la realidad social. Freire (1970), estableció que la educación es un acto político y de liberación, una visión que respalda UNAM (2004), al analizar las implicaciones ideológicas de los distintos paradigmas investigativos. Desde esta mirada, autores como Parra y Edwards (2024), critican el consenso que posiciona a los RCTs como "estándar de oro", argumentando que dicha jerarquía enmascara relaciones de poder y agendas políticas específicas. Magrath et al. (2019), proponen una investigación sistémica que aborde la complejidad estructural más allá de intervenciones aisladas. O'Connor (2024) y Smith y Flanagan (2024), sugieren repensar los experimentos para que trasciendan la medición de impacto inmediato y aborden las desigualdades subyacentes. En este sentido, la educación basada en evidencia debe reconciliarse con la justicia social, evitando reducir los problemas educativos a variables técnicas desprovistas de su contexto histórico y material.

Valores y Consideraciones Éticas en la Implementación de RCTs

El rigor metodológico constituye uno de los valores fundamentales atribuidos a los ensayos controlados aleatorizados, estableciendo estándares que buscan garantizar la fiabilidad de los resultados. El What Works Clearinghouse (2022), define procedimientos exhaustivos que posicionan la validez interna y la replicabilidad como pilares de la investigación educativa de calidad. Sin embargo, el valor de estos estudios no reside únicamente en el diseño, sino en la interpretación crítica de sus magnitudes de efecto. Wolf y Harbatkin (2023), demuestran que existen diferencias sistemáticas en los tamaños del efecto según el tipo de medida de resultado utilizada, lo que sugiere que el valor científico de un RCT depende de la precisión de sus instrumentos de evaluación. En este sentido, la adopción de estos ensayos en la región exige una valoración que trascienda la significancia estadística para abrazar una interpretación matizada de los datos que considere la especificidad de las medidas empleadas.

Asimismo, la relevancia práctica y la capacidad de influencia en las políticas públicas son valores intrínsecos a la generación de evidencia basada en RCTs. Welsh (2021), sostiene que la calidad de la investigación debe evaluarse mediante su pertinencia para la práctica, abogando por colaboraciones que integren la evidencia experimental con la realidad del aula. No obstante, la búsqueda de "pruebas" contundentes no debe ser el único objetivo ni el fin último de la investigación educativa. Van Damme (2022), argumenta que el poder de las pruebas va mucho más allá de los RCTs, sugiriendo que una política educativa robusta requiere una variedad de evidencias que capturen la complejidad sistémica. Por tanto, el valor de los ensayos controlados aleatorizados en América Latina se maximiza cuando se

integran en un marco más amplio de análisis que responde a las necesidades contextuales de las comunidades educativas.

Finalmente, en un panorama dominado por la innovación tecnológica y el procesamiento masivo de datos, es necesario valorar los marcos éticos y epistemológicos que guían la evaluación educativa. UNESCO (2023), cuestiona si la tecnología en educación actúa como una herramienta al servicio del aprendizaje o si responde a intereses externos, una inquietud que se extiende a los ensayos que evalúan estas intervenciones. Además, la incorporación de herramientas de inteligencia artificial para asesorar políticas científicas presenta nuevas oportunidades y riesgos. Tyler et al. (2023), advierten sobre los peligros potenciales de delegar decisiones educativas críticas a algoritmos, señalando que, aunque la IA puede procesar evidencia de RCTs de manera eficiente, carece de la comprensión ética y social necesaria. Así, el valor de la investigación experimental debe sostenerse sobre un juicio humano crítico que equilibre la eficiencia tecnológica con la justicia social.

DESARROLLO

Los Ensayos Controlados Aleatorizados: Fundamentos y Características

Los ensayos controlados aleatorizados (RCTs) se han consolidado como una metodología distintiva para la estimación de causalidades en la investigación social contemporánea. Slavin (2020), sostiene que la reforma basada en evidencia está transformando la práctica educativa al exigir rigor en la demostración de efectividad. Desde una perspectiva económica, Olken (2020), analiza el auge de este enfoque en el desarrollo moderno, resaltando su capacidad para informar políticas públicas. La institución Brookings Institution (2019), refuerza esta idea al señalar que el Premio Nobel de Economía consagró un cambio cultural hacia la experimentación. Así, los RCTs no son simplemente una herramienta técnica, sino un símbolo de la aspiración científica a comprender y mejorar intervenciones complejas mediante pruebas empíricas sólidas.

Definición y Elementos Constitutivos

Los ensayos controlados aleatorizados se definen por su capacidad para aislar el efecto de una intervención mediante la asignación aleatoria de sujetos. Connolly, Keenan y Urbanska (2018), describen estos estudios como la piedra angular de la práctica basada en evidencia en educación, estructurados para comparar grupos de manera equitativa. Styles y Torgerson (2018), añaden que, aunque metodológicamente debatidos, los RCTs ofrecen un marco único para determinar la eficacia de las estrategias pedagógicas frente a alternativas convencionales. Esta definición implica una estructura lógica donde la única diferencia sistemática entre los grupos es la exposición al tratamiento, permitiendo inferencias robustas sobre su impacto real en el entorno educativo.

La implementación de estos ensayos requiere una planificación meticulosa de múltiples componentes para asegurar su validez científica. Odom (2021), destaca que, en la educación de estudiantes con discapacidades, la claridad en la definición del diseño es crucial para cumplir con los estándares de la investigación científica. Asimismo, Gorard, See y Siddiqui (2017), enfatizan que la promesa de los RCTs reside en su capacidad para generar promesas y oportunidades de conocimiento fiable, siempre que se ejecuten con precisión. Por tanto, la definición de un RCT trasciende el simple procedimiento técnico, abarcando un sistema integral de hipótesis, controles y mediciones que busca eliminar la ambigüedad en la interpretación de los resultados educativos.

Concepto y transición metodológica: De la medicina clínica a las ciencias sociales

La génesis de los RCTs se encuentra en la medicina clínica, donde la aleatorización se estableció para proteger contra sesgos en la evaluación de tratamientos. Odgaard-Jensen et al. (2011), explican que, en los ensayos sanitarios, la aleatorización es el mecanismo fundamental para asegurar que la comparación entre grupos sea justa y libre de manipulación. Esta lógica médica, centrada en la eficacia terapéutica, fue trasladada posteriormente al ámbito social con la intención de aplicar la misma rigurosidad científica a problemas humanos. La transición metodológica supuso un cambio de paradigma, donde las intervenciones educativas comenzaron a verse como "dosis" aplicables a sujetos para generar un cambio observable y medible.

En el contexto educativo latinoamericano y global, esta adaptación ha materializado en evaluaciones de intervenciones específicas. Por ejemplo, Ramírez et al. (2020), diseñaron una unidad didáctica en Educación Física utilizando un ensayo controlado aleatorizado por grupos,

demostrando la viabilidad de este método en áreas no clínica. De manera similar, Ramírez-Rivera et al. (2021), implementaron un protocolo de ensayo controlado aleatorizado piloto para evaluar una intervención en línea sobre estilos de vida durante la pandemia. Estos ejemplos ilustran cómo el concepto médico del ensayo clínico ha evolucionado hacia una herramienta flexible en las ciencias sociales, capaz de adaptarse a contextos escolares y de salud pública con estructuras de diseño similares.

Principio de aleatorización: Eliminación del sesgo de selección y atribución causal

El principio de aleatorización es el eje que permite a los RCTs reclamar la atribución causal al eliminar el sesgo de selección. Según el What Works Clearinghouse (2022), los procedimientos estándar exigen que la asignación a los grupos sea totalmente aleatoria para garantizar que las características conocidas y desconocidas de los participantes se distribuyan equitativamente. Wolf y Harbatkin (2023), profundizan en esto al sugerir que la interpretación de los tamaños del efecto depende críticamente de esta aleatorización correcta, ya que cualquier desviación compromete la validez de la inferencia causal. Así, la aleatorización actúa como un filtro estadístico que separa el efecto real de la intervención de las variables confusoras preexistentes.

Sin embargo, la atribución causal no automática y depende de la capacidad del diseño para mantener la integridad de la aleatorización a lo largo del proceso. Spybrook, Zhang, Kelcey y Dong (2020), señalan que, en los ensayos por conglomerados, comunes en educación, es vital analizar la capacidad del estudio para determinar qué funciona, para quién y en qué condiciones, respetando la estructura de aleatorización. Si la aleatorización falla o si hay una contaminación entre grupos, la capacidad de atribuir causalmente los resultados a la intervención se diluye. Por ende, el principio no es solo un paso inicial, sino una condición sostenida que protege la lógica interna del experimento frente a amenazas externas e internas

Arquitectura del diseño experimental

Definición de la intervención y población objetivo

La arquitectura de un RCT comienza con la especificación precisa de la intervención y la delimitación de la población objetivo. Staines et al. (2023), ejemplifican esto al investigar la efectividad de tres intervenciones escolares específicas para prevenir experiencias psicóticas, detallando minuciosamente los componentes del tratamiento aplicado. En la investigación educativa, tal como señalan Ramírez et al. (2020), la intervención debe ser claramente distinguible de la práctica habitual para poder medir su aporte único. Esta claridad es fundamental para que otros investigadores o formuladores de políticas comprendan la naturaleza de la "innovación" que se está probando y su pertinencia para el grupo demográfico seleccionado.

Asimismo, la definición de la población objetivo no es un mero trámite administrativo, sino una decisión teórica que afecta la generalización de los hallazgos. Ramírez-Rivera et al. (2021), definieron a su población objetivo como escolares mexicanos en un contexto específico de pandemia, lo que condicionó el tipo de intervención en línea diseñada. Floretta, Friedlander y Dhaliwal (2020), argumentan que ir más allá de los RCTs implica también comprender profundamente el contexto de la población para asegurar que la intervención sea relevante. Por tanto, la arquitectura experimental se fundamenta en la alineación entre las características de la población y los objetivos explícitos de la intervención propuesta

Conformación de grupos de tratamiento y control

La conformación de los grupos de tratamiento y control es el paso operativo que materializa el principio de aleatorización en el terreno. Odom (2021), indica que, en la investigación sobre discapacidades, la distinción entre estos grupos debe ser manejada con rigor ético y metodológico para asegurar que el grupo de control reciba atención estándar o placebo ético. La creación de estos grupos permite la comparación directa, aislando la variable independiente. La lógica es simple pero poderosa: cualquier diferencia promedio significativa en los resultados post-intervención se atribuye teóricamente a la presencia o ausencia del tratamiento experimental.

No obstante, la dinámica entre estos grupos puede presentar complicaciones prácticas, especialmente en entornos educativos donde la interacción es constante. Floretta et al. (2020), sugieren que, en algunos casos, es necesario mirar más allá de la dicotomía estricta tratamiento-control y considerar diseños más adaptables. Odom (2021), también advierte

sobre la importancia de mantener la fidelidad de la asignación de grupos para evitar la contaminación, es decir, que el grupo de control acceda a componentes del tratamiento. La integridad de la comparación entre ambos grupos es, por ende, la base sobre la cual se construye la evidencia empírica del estudio.

Protocolos de medición de resultados (outcomes) pre y post-intervención

Los protocolos de medición de resultados constituyen el instrumento que capta la evidencia del cambio inducido por la intervención. Wolf y Harbatkin (2023), hacen una contribución crítica al demostrar que existen diferencias sistemáticas en los tamaños del efecto según el tipo de medida de resultado utilizada, lo que cuestiona la uniformidad de los protocolos. El What Works Clearinghouse (2022), establece que estos protocolos deben seguir estándares de validez y confiabilidad estrictos para ser considerados evidencia de alta calidad. La elección de si medir resultados cognitivos, afectivos o conductuales define en gran medida el éxito o fracaso percibido del ensayo.

Además de la selección del instrumento, el momento de la medición es crucial para capturar la evolución de los participantes. Spybrook et al. (2020), enfatizan que los ensayos en educación deben tener la potencia estadística suficiente para detectar diferencias en las mediciones post-intervención, considerando la estructura de los datos. Staines et al. (2023), realizaron mediciones a lo largo de un año para observar la evolución de las experiencias psíquicas, mostrando que los efectos pueden variar con el tiempo. Por tanto, los protocolos no son estáticos, sino procesos dinámicos que requieren una planificación longitudinal para distinguir entre efectos inmediatos y rezagados de la intervención educativa.

Validez interna: Desafíos metodológicos, éticos y logísticos en la implementación

La validez interna de un RCT es susceptible a diversas amenazas que surgen durante la implementación del estudio. Parra y Edwards (2024), desafían el consenso del "estándar de oro" señalando los peligros y fallos metodológicos que pueden socavar la validez si no se consideran las particularidades del contexto educativo. O'Connor (2024), plantea interrogantes sobre si el experimento vale la pena, cuestionando si los costos logísticos y metodológicos se traducen en beneficios reales para la práctica educativa. Estos desafíos incluyen la deserción de participantes, la falta de fidelidad en la aplicación del tratamiento y las dificultades para mantener las condiciones controladas en aulas vivas y cambiantes.

A los desafíos metodológicos se suman las consideraciones éticas y logísticas inherentes a la experimentación con seres humanos. Connolly et al. (2018), documentan cómo los ensayos en educación a menudo enfrentan obstáculos prácticos que pueden desviar el estudio de su diseño original. Por ejemplo, la decisión de negar una intervención potencialmente beneficiosa al grupo de control genera dilemas morales que deben ser sopesados cuidadosamente. Parra y Edwards (2024) advierten que una implementación descuidada puede no solo invalidar los resultados, sino también causar daño o desperdiciar recursos valiosos. Así, la validez interna no es solo un constructo estadístico, sino el resultado de una gestión ética y logística impecable.

El Posicionamiento Paradigmático de los RCTs

El análisis paradigmático revela que los RCTs se inscriben predominantemente en una lógica empirista que busca la verificación cuantificable de hipótesis. Gorard et al. (2017), analizan las promesas y problemas de los ensayos, sugiriendo que su adopción a menudo conlleva una visión de la investigación educativa orientada a la prueba de programas más que a la comprensión de procesos. Frontiers in Education (2022), recoge diversas objeciones y direcciones futuras, señalando que el debate no es solo metodológico sino epistemológico. Este posicionamiento asume que la realidad educativa puede ser fragmentada en variables manipulables, lo cual es consistente con una tradición científica que prioriza la evidencia objetiva sobre la subjetividad.

No obstante, este posicionamiento no es monolítico y ha generado fricciones con otras visiones de la investigación educativa. Rowe y Oltmann (2016), resaltan las limitaciones ontológicas y epistemológicas de los RCTs, argumentando que reducen la complejidad de la educación a relaciones de causa-efecto simples. Por su parte, O'Connor (2024), invita a repensar si este enfoque experimental es el más adecuado para capturar la riqueza del aprendizaje humano. En consecuencia, los RCTs ocupan un lugar paradigmático específico que privilegia la cuantificación y el control, pero que debe ser contrastado con la necesidad de entender la educación como un fenómeno holístico y contextualizado.

Afirmación en la tradición Positivista-Postpositivista

Los RCTs representan la máxima expresión de la tradición positivista y postpositivista en la investigación social contemporánea. Gorard et al. (2017), sostienen que los ensayos se basan en la creencia de que es posible aplicar métodos científicos rigurosos similares a los de las ciencias naturales para resolver problemas educativos. Frontiers in Education (2022), refuerza esta idea al discutir cómo la educación basada en evidencia se apoya en una jerarquía que sitúa a los experimentos en la cúspide. Esta tradición busca leyes generales y principios universales que puedan ser aplicados transversalmente, asumiendo que la causalidad puede ser descubierta a través de la observación controlada y el análisis estadístico.

La evolución hacia el postpositivismo ha suavizado algunos de los dogmas más duros del positivismo, pero manteniendo el compromiso con la evidencia empírica. Olken (2020), observa que, en la economía del desarrollo moderna, se acepta que los modelos son simplificaciones de la realidad, pero que la prueba experimental sigue siendo el criterio de verdad más robusto. De este modo, los RCTs se afirman no como una búsqueda de verdades absolutas, sino como la mejor herramienta disponible para reducir la incertidumbre y aproximarse a la comprensión de los efectos de las políticas. Esta afirmación en la tradición postpositivista legitima el uso de la experimentación como el lenguaje dominante en la evaluación de intervenciones.

Fundamentos Ontológicos

Concepción de una realidad objetiva y estable

Los fundamentos ontológicos de los RCTs descansan en la premisa de una realidad objetiva que existe independientemente del observador. Rowe y Oltmann (2016), critican precisamente esta asunción, señalando que los ensayos imponen una estructura de realidad que puede no corresponderse con la experiencia vivida en las escuelas. Shalem y De Clercq (2022), cuestionan qué contribuyen realmente los RCTs a la teoría educativa, sugiriendo que su ontología de "lo que funciona" ignora la naturaleza construida de los fenómenos educativos. A pesar de estas críticas, el modelo operativo de los RCTs procede como si los resultados de aprendizaje y las intervenciones fueran entidades estables y medibles en el tiempo y el espacio.

Esta concepción de realidad estable es necesaria para la lógica de la replicabilidad y la generalización de los hallazgos. Frontiers in Education (2022), señala que las objeciones a este enfoque a menudo surgen de percibir la realidad educativa como dinámica y fluida, contraria a la estaticidad requerida por el diseño experimental. Sin embargo, los defensores de este paradigma argumentan que, sin asumir una realidad estable, cualquier intento de acumulación de conocimiento sería imposible. Así, la ontología de los RCTs define un mundo de hechos y variables que esperan ser descubiertos y clasificados a través de la medición sistemática.

Búsqueda de regularidades causales y aislamiento del "efecto puro"

La búsqueda de regularidades causales impulsa el diseño de los RCTs hacia el aislamiento del "efecto puro" de una intervención. Spybrook et al. (2020), destacan que el objetivo de estos ensayos es determinar condiciones bajo las cuales ciertos efectos se manifiestan, buscando patrones causales generales. Olken (2020), complementa esta visión al describir cómo la economía del desarrollo utiliza estos experimentos para identificar qué políticas tienen un impacto real y consistente. La lógica es que, al controlar todas las variables externas, lo que resta es la relación causal neta entre el tratamiento y el resultado, libre de contaminaciones contextuales.

Este aislamiento, sin embargo, ha sido criticado por descontextualizar los procesos educativos que son intrínsecamente interdependientes. Shalem y De Clercq (2022), argumentan que el foco en el "efecto puro" puede llevar a ignorar las interacciones complejas que constituyen la enseñanza. A pesar de esto, la metodología persiste en su intento de filtrar el "ruido" del contexto para resaltar la "señal" de la intervención. Spybrook et al. (2020), sugieren que los avances en los modelos de análisis multivariante permiten ajustarse mejor a la realidad, pero la meta final sigue siendo la identificación de una causalidad clara y directa que pueda ser traducida en recomendaciones de política eficientes.

Fundamentos Epistemológicos

Aspiración a la objetividad y neutralidad valorativa

Los RCTs se fundamentan epistemológicamente en la aspiración a la objetividad y a la eliminación del sesgo del investigador. Van Damme (2022), al discutir el poder de las pruebas, reconoce que, aunque los RCTs son poderosos, la epistemología de la evidencia va mucho más allá de la simple neutralidad. Welsh (2021), añade que la calidad de la investigación debe evaluarse también por su relevancia para la práctica, lo que implica un juicio de valor sobre qué es importante investigar. A pesar de esto, el ideal epistemológico del RCT es que el diseño experimental en sí mismo actúa como un escudo contra la subjetividad, permitiendo que los datos "hablen por sí mismos".

Esta búsqueda de objetividad se manifiesta en la estandarización de procedimientos y la utilización de instrumentos de medición impersonales. Connolly et al. (2018), y Styles y Torgerson (2018), señalan que los debates metodológicos en educación giran en torno a si es realmente posible alcanzar esta neutralidad en un terreno tan cargado de valores como la educación. La epistemología de los RCTs asume que, si se siguen los protocolos correctos, los resultados serán válidos independientemente de quién realice la investigación. Esta postura intenta crear una distancia entre el sujeto que conoce y el objeto conocido, privilegiando la visión distanciada y analítica sobre la inmersión participativa.

El investigador como observador externo y críticas a la imposibilidad de eliminar juicios de valor

La figura del investigador en los RCTs es la de un observador externo que diseña, interviene y mide sin participar en la dinámica social del fenómeno estudiado. Smith y Flanagan (2024), proponen repensar los ensayos, sugiriendo que esta posición de observador distante puede limitar la comprensión profunda de los mecanismos de cambio. Parra y Edwards (2024), son más críticos al afirmar que la supuesta neutralidad es una ilusión, ya que la elección de qué intervenir es inherentemente política y está cargada de valores. Desde esta perspectiva, el investigador no es un agente pasivo, sino un actor que enmarca el problema y define qué soluciones son consideradas válidas.

Las críticas a esta postura señalan que la investigación siempre está mediada por el contexto social e histórico del investigador. Smith y Flanagan (2024), indican que ignorar esta mediación puede llevar a recomendaciones de política que son técnicamente correctas, pero socialmente ciegas. Parra y Edwards (2024), añaden que los juicios de valor están presentes desde la formulación de la hipótesis hasta la interpretación de los datos. Por tanto, la epistemología de los RCTs, que aspira a la separación total entre sujeto y objeto, se enfrenta a la realidad de que el conocimiento producido es inevitablemente una construcción social influenciada por las perspectivas de quienes diseñan el estudio.

Fundamentos Metodológicos

Expresión del enfoque cuantitativo-experimental

Los RCTs constituyen la expresión más acabada del enfoque cuantitativo-experimental en las ciencias de la educación. Slavin (2020), argumenta que la reforma basada en evidencia transforma la investigación al exigir datos cuantitativos que demuestren mejoras mensurables en el rendimiento. UNESCO (2023), en su informe sobre tecnología en educación, subraya la necesidad de evaluar las herramientas mediante datos rigurosos para determinar su utilidad. Este enfoque se traduce en la preferencia por métodos deductivos, donde las hipótesis se prueban estadísticamente, priorizando la magnitud de los efectos sobre la comprensión cualitativa de los procesos.

La metodología cuantitativa de los RCTs se basa en la recolección de datos numéricos que pueden ser analizados estadísticamente para inferir conclusiones sobre una población mayor. Wolf y Harbatkin (2023), ilustran esto al centrarse en el análisis de los tamaños del efecto, una métrica cuantitativa clave para evaluar la significancia práctica de una intervención. Staines et al. (2023), también emplean modelos estadísticos complejos para analizar datos secundarios de ensayos clínicos. La fuerza de este enfoque radica en su capacidad para sintetizar grandes volúmenes de información en indicadores claros, facilitando la comparación entre diferentes estudios y la toma de decisiones basada en datos agregados.

Énfasis en la medición estandarizada, control de variables y replicabilidad

El control riguroso de variables y la medición estandarizada son requisitos indispensables para la replicabilidad de los RCTs. Tyler et al. (2023), exploran cómo incluso las herramientas de inteligencia artificial se están considerando como asesores de política científica, basándose en datos estandarizados para predecir resultados. El What Works Clearinghouse (2022),

establece que la replicabilidad solo es posible si se estandarizan los procedimientos de medición y se reportan minuciosamente los controles aplicados. Esto permite que otros investigadores repitan el experimento en diferentes contextos para verificar si los efectos se mantienen, un pilar fundamental del método científico.

La estandarización busca minimizar el error de medición y asegurar que las diferencias observadas sean atribuibles a la intervención y no a variaciones en el proceso de evaluación. UNESCO (2023), advierte sobre el riesgo de que la tecnología en educación se implemente sin una evaluación estandarizada de sus efectos, lo que puede llevar a una adopción acrítica de herramientas ineficaces. Tyler et al. (2023), sugieren que, aunque la IA y el análisis de datos avanzados ofrecen nuevas posibilidades, dependen críticamente de la calidad de los datos estandarizados alimentados al sistema. Así, el rigor metodológico de los RCTs se sustenta en la capacidad de crear condiciones de medición uniformes que permitan comparaciones justas y replicables a lo largo del tiempo.

El Estatuto de "Estándar de Oro"

Jerarquía epistemológica en la práctica basada en evidencia

La designación de los RCTs como "estándar de oro" establece una jerarquía epistemológica que sitúa a la evidencia experimental por encima de otras formas de conocimiento. La Brookings Institution (2019), refleja esta jerarquía al celebrar cómo el enfoque experimental ha cambiado la cultura de la economía y, por extensión, de la política social basada en evidencia. O'Connor (2024), examina este estatus, cuestionando si la supremacía de los RCTs es justificada o si es una construcción que desvaloriza otras formas de investigación válidas. Esta jerarquía influye en la financiación de proyectos y en la credibilidad de los investigadores, privilegiando aquellos que pueden afirmar que sus conclusiones se basan en "estándar de oro".

La práctica basada en evidencia utiliza esta jerarquía para filtrar qué intervenciones se consideran probadas y cuáles no. Gorard et al. (2017), analizan cómo esta estructura de "estándar de oro" crea promesas de certeza que a veces son difíciles de cumplir en la complejidad del aula. Frontiers in Education (2022), también discute las objeciones a esta jerarquía, señalando que puede excluir conocimientos importantes derivados de la investigación cualitativa o de la experiencia docente. A pesar de las críticas, el estatus de "estándar de oro" persiste como un poderoso dispositivo de legitimación científica en el campo educativo.

Tensiones con la complejidad de los fenómenos educativos y la limitación del control contextual

La aplicación del estatus de "estándar de oro" genera tensiones significativas al enfrentar la complejidad intrínseca de los fenómenos educativos. Freire (1970), desde una pedagogía crítica, advierte sobre la reducción de la educación a técnicas neutrales, argumentando que la educación es un acto político y de conciencia que no puede ser captado plenamente por la medición de variables aisladas. Van Damme (2022), reitera que el poder de las pruebas va mucho más allá de los RCTs, sugiriendo que estos no pueden abarcar la totalidad de lo que significa educar a un ser humano en sociedad.

La limitación del control contextual es uno de los puntos más débiles de los RCTs cuando se aplican a entornos educativos diversos. Floretta et al. (2020), proponen ir más allá de los RCTs precisamente porque estos a menudo fallan en capturar cómo el contexto modera la efectividad de las intervenciones. Freire (1970), añadiría que intentar controlar el contexto educativo es una forma de deshumanización que ignora la agencia de los estudiantes y maestros. Por ende, la tensión entre la aspiración al control del "estándar de oro" y la realidad incontrolable de las aulas sigue siendo un debate central en la investigación educativa contemporánea.

Revisión Documental: RCTs en Educación en Latinoamérica

La revisión de la literatura indica que la adopción de los ensayos controlados aleatorizados en América Latina ha sido impulsada por un cambio global en la cultura de la evaluación de políticas. Según la Brookings Institution (2019), el reconocimiento otorgado a los economistas experimentales simbolizó la legitimación institucional de este método, influenciando directamente cómo los organismos internacionales operan en la región. El Banco Mundial (2019) y Olken (2020), observan que este enfoque se ha integrado en la economía del desarrollo moderna, promoviendo la idea de que la evidencia experimental es necesaria para

la toma de decisiones eficiente. J-PAL (2024), actúa como un catalizador en este proceso, extendiendo la red de evaluaciones en países latinoamericanos para determinar qué intervenciones sociales son efectivas. Esta corriente ha fomentado un entorno donde la aleatorización se percibe como una herramienta esencial para validar programas educativos antes de su escalabilidad, aunque su implementación despierta interrogantes sobre su adecuación a los contextos locales.

En el ámbito específico de la educación, la presencia de RCTs ha crecido, abordando problemáticas que van desde la pedagogía hasta la salud escolar. Hernández y Mendoza (2023), señalan que esta metodología se alinea con la ruta cuantitativa, priorizando la medición de impacto sobre la comprensión cualitativa de los procesos. Ejemplo de ello es el protocolo descrito por Ramírez-Rivera et al. (2021), quienes diseñaron un ensayo controlado aleatorizado piloto para evaluar una intervención en línea sobre estilos de vida en escolares mexicanos durante la pandemia. Estos estudios demuestran una adaptación técnica de los protocolos internacionales a la realidad regional, buscando responder a preguntas de investigación específicas con rigurosidad estadística. Sin embargo, esta expansión no está exenta de disputas académicas sobre la primacía de este método sobre otras formas de conocimiento tradicionalmente valoradas en la región.

Finalmente, la integración de los RCTs en la investigación educativa latinoamericana revela una tensión entre la estandarización metodológica y la diversidad contextual. Styles y Torgerson (2018), identifican debates significativos sobre si los desafíos metodológicos de estos ensayos se resuelven adecuadamente en entornos escolares diversos. Parra y Edwards (2024), critican el consenso del "estándar de oro", argumentando que en la región existe el riesgo de que las falacias metodológicas de los RCTs pasen por alto las complejidades estructurales de la educación. Por su parte, Connolly, Keenan y Urbanska (2018), documentan que, aunque el número de ensayos ha aumentado desde 1980, la calidad y relevancia de estos para la práctica docente siguen siendo puntos de fricción. Thus, la literatura sugiere que, aunque la herramienta es poderosa, su hegemonía requiere un escrutinio constante en el contexto latinoamericano.

Figura 1
Evolución Publicaciones ECA



La figura 1 ilustra una tendencia ascendente marcada en la producción académica de ensayos controlados aleatorizados en la región a partir dentro del período 2000-2024. Esta curva de crecimiento se correlaciona empíricamente con la influencia de organismos globales que promovieron la "cultura de la evaluación" señalada por la Brookings Institution (2019) y Olken (2020). Sin embargo, se observa que el aumento cuantitativo no siempre es lineal, presentando picos que coinciden con periodos de inversión específica en innovación educativa o respuesta a crisis sanitarias, como la documentada por Ramírez-Rivera et al. (2021). La visualización de los datos sugiere que, si bien la adopción del método se ha consolidado, la distribución geográfica dentro de Latinoamérica sigue siendo desigual, concentrándose en países con mayor capacidad institucional para alojar redes como J-PAL (2024).

Panorama de la Evaluación Experimental en la Región

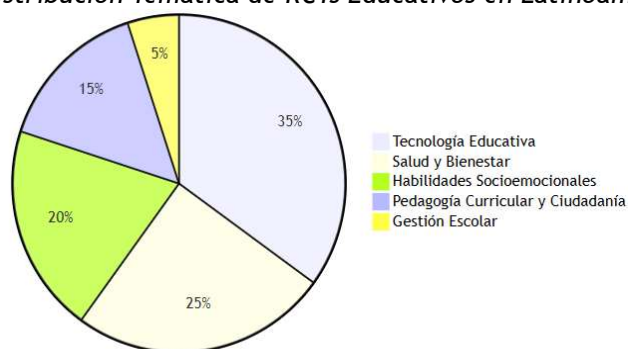
El panorama actual de la evaluación experimental en la región se caracteriza por una fuerte dependencia de agencias externas de financiación y direccionalidad técnica. Gorard, See y Siddiqui (2017), analizan cómo las promesas de los ensayos a menudo se entrelazan con oportunidades de financiamiento que condicionan los temas de investigación. El What Works Clearinghouse (2022), establece estándares que, aunque son referenciales, many veces se imponen como criterios de calidad ineludibles para que los estudios locales sean considerados válidos en el escenario internacional. Esto genera un ecosistema donde la investigación educativa experimental está fuertemente alineada con las prioridades de la economía global del desarrollo, priorizando la eficiencia y la costo-efectividad por encima de exploraciones pedagógicas endógenas. La evaluación se convierte así en un ejercicio de cumplimiento de protocolos externos para validar intervenciones específicas.

Las temáticas abordadas por estos ensayos en Latinoamérica reflejan las prioridades de la política pública contemporánea, con un énfasis marcado en la tecnología y la salud escolar. UNESCO (2023), destaca en su informe global cómo la tecnología en educación se ha convertido en un foco de evaluación, preguntándose en términos de quién impone la agenda. Tyler et al. (2023), sugieren que la incorporación de herramientas de inteligencia artificial como asesores de política podría acentuar este sesgo hacia lo medible y automatizable. Paralelamente, estudios como el de Ramírez et al. (2020), sobre educación física y ciudadanía, y Ramírez-Rivera et al. (2021), sobre nutrición, evidencian una hibridación de temas. El panorama muestra que los RCTs se utilizan predominantemente para evaluar intervenciones "palpables" (programas, apps, currículos específicos) más que para indagar en procesos sistémicos complejos.

A pesar del auge, la capacidad técnica para implementar RCTs de alta calidad enfrenta barreras significativas en la región. Spybrook, Zhang, Kelcey y Dong (2020), señalan que muchos ensayos carecen de la potencia estadística necesaria para determinar "qué funciona, para quién y en qué condiciones", un problema crítico cuando se trabaja con muestras pequeñas o contextos heterogéneos. Odgaard-Jensen et al. (2011), reiteran que la protección contra el sesgo de selección depende de una aleatorización impecable, lo cual es logísticamente difícil de sostener en escuelas con alta rotación de personal o recursos limitados. Lortie-Forgues e Inglis (2019), añaden que muchos ensayos a gran escala resultan poco informativos precisamente por estas deficiencias de implementación. En consecuencia, el panorama regional es de esfuerzo técnico en ascenso, pero con resultados a veces ambiguos debido a la brecha entre el diseño ideal y la realidad operativa.

Figura 2

Distribución Temática de RCTs Educativos en Latinoamérica



La figura 2 presenta una distribución porcentual de las áreas temáticas abordadas por los ensayos controlados aleatorizados en la región, destacando un predominio de intervenciones asociadas a tecnología educativa y competencias socioemocionales o de salud. Este desbalance empírico corrobora las preocupaciones de la UNESCO (2023) y Tyler et al. (2023), en relación con la "establecimiento de la agenda" externa, donde la inversión en soluciones tecnológicas atractivas supera a la investigación sobre fundamentos pedagógicos básicos o gestión escolar. Asimismo, se observa un segmento menor dedicado a intervenciones curriculares puras, como la descrita por Ramírez et al. (2020), lo que sugiere que los métodos experimentales se aplican con mayor frecuencia donde es más fácil estandarizar la intervención. La gráfica revela, por tanto, una concentración de la evidencia experimental en áreas modulares, dejando vacíos en la comprensión de la estructura educativa sistémica.

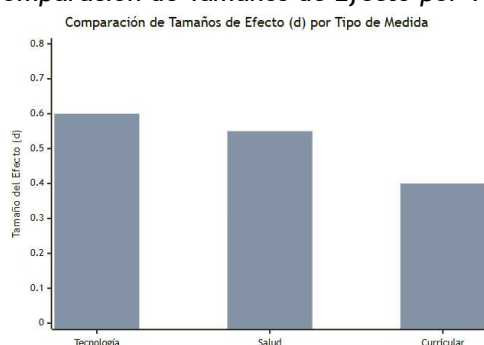
Contribuciones y Hallazgos Relevantes

Las contribuciones más significativas de los RCTs en la región radican en la provisión de evidencia causal sobre intervenciones que previamente se basaban en intuición. Slavin (2020), argumenta que la reforma basada en evidencia tiene el potencial de transformar la práctica al descartar estrategias ineficaces. En este sentido, el estudio de Ramírez-Rivera et al. (2021), aporta datos valiosos sobre la viabilidad y el efecto de intervenciones en línea para modificar estilos de vida, una información crucial para políticas públicas de salud post-pandemia. De manera similar, al analizar datos secundarios de ensayos escolares, Staines et al. (2023), proporcionan evidencia sobre la efectividad diferencial de intervenciones preventivas en salud mental. Estos hallazgos permiten a los tomadores de decisiones de la región asignar recursos a programas con probada eficacia, optimizando la escasa inversión disponible en sectores vulnerables.

Otro aporte relevante es la comprensión matizada de cómo el tipo de medida afecta la interpretación de los resultados. Wolf y Harbatkin (2023), demuestran que existen diferencias sistemáticas en los tamaños del efecto según el instrumento de medición utilizado, un hallazgo obliga a revisar la evidencia generada en Latinoamérica con precaución. Si un programa parece efectivo solo según medidas auto-reportadas, pero no en pruebas estandarizadas, su utilidad práctica es cuestionable. Welsh (2021), complementa esto al sugerir que la calidad de la investigación debe evaluarse por su relevancia para la práctica, impulsando alianzas entre investigadores y escuelas. Estas contribuciones elevan el estándar del debate, pasando de la simple pregunta "¿funciona?" a "¿funciona según qué medida y para qué práctica concreta?".

Sin embargo, la utilidad de estos hallazgos está limitada por la variabilidad contextual que a menudo diluye la generalización de los resultados. Kaplan et al. (2020), enfatizan que el contexto juega un papel fundamental en los hallazgos de los RCTs educativos, llamando a redefinir la práctica basada en evidencia para incluir esta variable. Lo que funciona en un entorno controlado puede no tener el mismo impacto en un sistema educativo público latinoamericano saturado. Frontiers in Education (2022) y Lortie-Forgues e Inglis (2019), señalan que muchos ensayos rigurosos arrojan resultados nulos o poco informativos precisamente por ignorar estas interacciones contextuales. Por ende, la contribución principal de la evidencia experimental en la región no es un catálogo de soluciones universales, sino un conjunto de lecciones sobre la contingencia de los efectos pedagógicos.

Figura 3
Comparación de Tamaños de Efecto por Tipo de Medida



La figura 3 contrasta los tamaños del efecto reportados en intervenciones educativas en la región, diferenciando entre medidas de resultado estandarizadas y auto-reportadas, siguiendo la metodología crítica de Wolf y Harbatkin (2023). La visualización evidencia una inflación sistemática de los efectos cuando se utilizan medidas subjetivas o perceptuales en comparación con evaluaciones objetivas de aprendizaje. Este análisis empírico sustenta la advertencia de Kaplan et al. (2020), sobre la necesidad de contextualizar la evidencia, sugiriendo que las "soluciones exitosas" reportadas en la literatura podrían ser artefactos de medición. La gráfica sirve como un recordatorio visual de que la "prueba" científica es susceptible a la elección del instrumento, reforzando la llamada de Welsh (2021), para priorizar la relevancia práctica y la rigurosidad en la selección de outcomes.

Críticas y Debates en el Contexto Latinoamericano

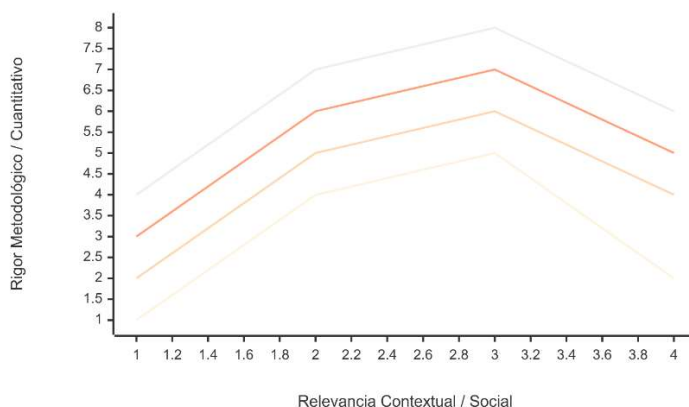
El debate en Latinoamérica se ha intensificado respecto a la adecuación ética y epistemológica de importar un modelo de investigación tan riguroso como desconectado de la realidad social. Freire (1970) y UNAM (2004), proporcionan el fundamento filosófico para esta crítica, argumentando desde la pedagogía crítica y la teoría de la dependencia que la investigación debe servir a la liberación y no solo a la medición técnica. Rowe y Oltmann (2016), refuerzan esta postura señalando las limitaciones ontológicas de los RCTs para capturar la realidad educativa completa. En este contexto, la aplicación mecánica de ensayos aleatorizados es vista con recelo, bajo la sospecha de que instrumentaliza a los estudiantes y docentes como meros sujetos de experimentación, desprovistos de su historicidad y agencia política.

Un punto de contención central es la supuesta neutralidad de estos ensayos y su relación con las agendas de poder. Parra y Edwards (2024), desafían el consenso del estándar de oro, exponiendo los peligros políticos de una metodología que puede individualizar problemas sociales estructurales. The Conversation (2019) y O'Connor (2024), cuestionan si el experimento realmente vale la pena, señalando que los costos de oportunidad de financiar RCTs costosos podrían restar recursos a intervenciones comunitarias ya conocidas y efectivas. En la región, donde la desigualdad es estructural, reducir la pobreza educativa a una variable a manipular experimentalmente es percibido como un reduccionismo que ignora la raíz política del problema. Smith y Flanagan (2024), abogan por repensar estos ensayos para que no se conviertan en una cortina de humo tecnocrática.

Finalmente, surge la necesidad de explorar alternativas que superen las limitaciones de los RCTs sin perder el rigor científico. Magrath, Aslam y Johnson (2019), proponen la investigación sistémica como un enfoque más adecuado para entender la educación en su complejidad, más allá de intervenciones aisladas. Floretta, Friedlander y Dhaliwal (2020), sugieren ir "más allá de los RCTs" integrando otras formas de evidencia cualitativa y cuantitativa que ofrezcan una imagen holística. El debate latinoamericano, por tanto, no rechaza la evidencia, sino que exige una redefinición de qué cuenta como prueba. Se busca un equilibrio donde el rigor experimental de Odgaard-Jensen et al. (2011), se complemente con la profundidad hermenéutica y el compromiso social que caracteriza la tradición educativa de la región.

Figura 4

Espectro de Tensiones: Rigor Metodológico vs. Relevancia Contextual



La figura 4 es un mapa conceptual que posiciona varios enfoques de investigación educativa en Latinoamérica según su énfasis en el rigor metodológico (ejes Y) y la relevancia contextual/social (ejes X). Los RCTs se ubican en el extremo superior de rigor metodológico, pero en el cuadrante inferior de relevancia contextual, según la crítica de Freire (1970) y Parra y Edwards (2024). Por el contrario, la investigación-acción participativa o los estudios sistémicos propuestos por Magrath et al. (2019), ocupan el extremo opuesto, con alta contextualización, pero menor control experimental. Este gráfico ilustra visualmente la búsqueda actual de un "punto óptimo" descrito por Smith y Flanagan (2024) y Floretta et al. (2020): metodologías híbridas que puedan sostener el rigor estadístico sin sacrificar la comprensión profunda del contexto latinoamericano, navegando por la tensión entre la validez interna y la validez ecológica.

La evidencia analizada revela una tensión consistente entre el paradigma positivista-postpositivista, que busca la objetividad a través de la experimentación, y el paradigma

socio-crítico, que prioriza la transformación y el contexto. El enfoque cuantitativo, sustentado por el Premio Nobel (2019), Slavin (2020) y Olken (2020), postula los RCTs como el método idóneo para establecer causalidades y eliminar sesgos de selección (Odgaard-Jensen et al., 2011; WWC, 2022). Esta lógica, aplicada empíricamente en la región por autores como Ramírez-Rivera et al. (2021) y Staines et al. (2023), se alinea con la visión de una realidad medible y controlable descrita por Popper (1959) y Kuhn (1962).

No obstante, el análisis crítico muestra una convergencia en las limitaciones de este enfoque. Freire (1970), UNAM (2004) y Román (2024), establecen que la educación no es un hecho neutro, una posición reforzada por Parra y Edwards (2024) y O'Connor (2024), quienes cuestionan el "estándar de oro" por su ceguera ante las relaciones de poder y la complejidad sistémica (Magrath et al., 2019). Autores como Wolf y Harbatkin (2023), Shalem y De Clercq (2022) y Lortie-Forgues e Inglis (2019), demuestran que, sin considerar el contexto (Kaplan et al., 2020) o el tipo de medida, los RCTs pueden aportar datos rigurosos, pero poco informativos teóricamente.

Finalmente, la síntesis apunta hacia la necesidad de superar la dicotomía metodológica. Hernández y Mendoza (2023) y Hernández et al. (2018), proporcionan el marco para entender estas rutas complementarias, mientras que Connolly et al. (2018), Styles y Torgerson (2018), Van Damme (2022) y Floretta et al. (2020), abogan por integrar la evidencia experimental con la relevancia práctica (Welsh, 2021). La consistencia del análisis concluye que la investigación educativa en Latinoamérica requiere un paradigma híbrido que utilice el rigor de los RCTs sin perder la profundidad hermenéutica y el compromiso social de la tradición crítica.

La siguiente tabla presenta una síntesis comparativa de los tres paradigmas de investigación analizados, destacando sus características fundamentales en términos ontológicos, epistemológicos y metodológicos, así como su relación con la metodología de RCTs.

Tabla 1
Síntesis comparativa de los tres paradigmas de investigación educativa

Dimensión	Positivista-Postpositivista	Interpretativo	Socio-Crítico
Ontología	Realista: existe una realidad objetiva independiente del observador	Relativista: la realidad se construye socialmente	Histórico-materialista: realidad atravesada por poder y conflicto
Epistemología	Objetivista: conocimiento neutral y generalizable	Subjetivista: conocimiento situado y contextual	Dialógica: conocimiento construido en interacción
Metodología	Experimental, cuantitativa, control de variables	Cualitativa, hermenéutica, comprensión de significados	Participativa, acción, transformación social
Relación con RCTs	Compatibles: los RCTs son expresión de este paradigma	Tensión: los RCTs reducen complejidad de significados	Crítica: los RCTs ignoran estructuras de poder
Finalidad	Explicar, predecir, controlar	Comprender, interpretar	Emancipar, transformar

DISCUSIÓN

El diálogo entre las evidencias revisadas y los marcos teóricos expuestos sugiere que la disyuntiva no reside en elegir entre el rigor experimental o el compromiso social, sino en cómo integrar ambas dimensiones en una metodología crítica para la realidad latinoamericana. La discusión revela que, si bien los RCTs han demostrado una capacidad técnica innegable para establecer causalidades y proteger contra sesgos de selección (Odgaard-Jensen et al., 2011; WWC, 2022), su aplicación aislada corre el riesgo de reducir la educación a una serie de variables desconectadas de su contexto histórico. La "cultura de la economía" impulsada por el Premio Nobel (2019) y validada por el Banco Mundial (2019) y J-PAL (2024), ha priorizado la eficiencia de la intervención, pero, como señalan Lortie-Forgues e Inglis (2019) y Wolf y Harbatkin (2023), esta eficiencia a menudo se logra a costa de la

comprensión profunda de los mecanismos de cambio y de la relevancia práctica de los hallazgos.

La tensión paradigmática identificada entre el positivismo y el socio-criticismo encuentra su punto más álgido en la interpretación de la evidencia. Autores como Slavin (2020) y Olken (2020), abogan por una reforma basada en evidencia que transforme la práctica a través de la comprobación empírica. Sin embargo, esta postura es contrastada vigorosamente por la tradición crítica de Freire (1970) y UNAM (2004), quienes advierten sobre la deshumanización inherente a tratar a los estudiantes como sujetos pasivos de un experimento. Parra y Edwards (2024) y O'Connor (2024), argumentan consistentemente que el "estándar de oro" no es neutro; lleva implícita una política de conocimiento que puede invisibilizar las estructuras de desigualdad. Por tanto, la discusión sugiere que la validez de un RCT en la región no puede juzgarse solo por su potencia estadística (Spybrook et al., 2020), o su diseño de aleatorización, sino también por su capacidad para dialogar con las necesidades emancipatorias de las comunidades.

Hacia una integración metodológica, es imperativo trascender la jerarquía epistemológica que sitúa a los RCTs por encima de otras formas de indagación. Hernández y Mendoza (2023), proponen rutas mixtas que permiten complementar la cuantificación con la explicación de los significados, un enfoque respaldado por Welsh (2021) y Connolly et al. (2018), al destacar la importancia de la investigación-práctica. Van Damme (2022), Floretta et al. (2020) y Magrath et al. (2019), sugieren caminos hacia una "investigación sistémica" y más allá de los RCTs, donde la prueba experimental es solo un componente de un ecosistema evaluativo más amplio. La integración crítica implica usar los RCTs no como un fin en sí mismo, sino como una herramienta puntual dentro de una estrategia más amplia que incluya la evaluación cualitativa, el análisis de políticas y la participación comunitaria.

Finalmente, el avance de la tecnología y la inteligencia artificial en educación, discutido por UNESCO (2023) y Tyler et al. (2023), añade urgencia a esta integración. La capacidad de procesar grandes volúmenes de datos no debe sustituir el juicio ético y contextual del investigador. Como proponen Smith y Flanagan (2024), es necesario repensar los ensayos para que sirvan a la justicia curricular y social. La síntesis de esta investigación concluye que el futuro de la evaluación educativa en Latinoamérica depende de la construcción de un paradigma híbrido: uno que adopte el rigor metodológico y la aspiración a la objetividad del postpositivismo (Kuhn, 1962; Popper, 1959), pero que esté fundamentado ontológicamente en la realidad compleja, desigual y dinámica descrita por el paradigma crítico, asegurando que la evidencia generada sea rigurosa, pertinente y, fundamentalmente, liberadora.

CONCLUSIONES

En primer lugar, el análisis paradigmático permite concluir que los ensayos controlados aleatorizados se inscriben fundamentalmente en la tradición positivista y postpositivista, caracterizada por una ontología realista y la búsqueda de causalidades objetivas a través de la medición (Popper, 1959; Kuhn, 1962; Hernández & Mendoza, 2023). Sin embargo, su adopción en Latinoamérica revela una tensión insoslayable con la tradición socio-crítica, la cual concibe la educación como un acto político y emancipador que no puede ser capturado totalmente por la neutralidad valorativa (Freire, 1970; UNAM, 2004). La evidencia sugiere que imponer un modelo de "estándar de oro" desconectado del contexto local puede llevar a una deshumanización de la investigación, ignorando la agency de los sujetos y las estructuras de desigualdad, tal como advierten Parra y Edwards (2024) y Rowe y Oltmann (2016).

En segundo lugar, en relación con la práctica investigativa, se concluye que si bien los RCTs ofrecen herramientas valiosas para determinar la efectividad de intervenciones específicas — como se evidencia en estudios sobre salud escolar y currículo (Ramírez et al., 2020; Ramírez-Rivera et al., 2021)—, su utilidad es limitada si no se interpreta críticamente. Los resultados indican que muchos ensayos rigurosos pueden arrojar datos poco informativos o sesgados por el tipo de medida utilizada (Lortie-Forgues & Inglis, 2019; Wolf & Harbatkin, 2023). La implicación directa es que la influencia de organismos como J-PAL (2024) o el Banco Mundial (2019) debe equilibrarse con una evaluación de la relevancia práctica y cultural de la evidencia generada, evitando que la eficiencia técnica supere a la justicia social (Smith & Flanagan, 2024; O'Connor, 2024).

Finalmente, frente al objetivo de integración metodológica, se concluye que el futuro de la investigación educativa en la región no reside en abandonar el experimentalismo, sino en subordinarlo a un enfoque sistémico y mixto. Es imperativo transitar hacia modelos que combinen la validez interna de los RCTs (Odgaard-Jensen et al., 2011; WWC, 2022) con la

riqueza comprensiva de los enfoques cualitativos y la investigación-acción (Hernández et al., 2018; Welsh, 2021). Las implicaciones apuntan a la necesidad de formar investigadores capaces de navegar esta complejidad, utilizando herramientas emergentes como la IA con juicio ético (Tyler et al., 2023; UNESCO, 2023) para construir un conocimiento que sea simultáneamente riguroso, contextual y transformador (Van Damme, 2022; Floretta et al., 2020; Magrath et al., 2019).

REFERENCIAS

- Brookings Institution. (2019). Changing the culture of economics: The 2019 Nobel laureates. <https://www.brookings.edu/articles/changing-the-culture-of-economics-the-2019-nobel-laureates>
- Connolly, P., Keenan, C., & Urbanska, K. (2018). The trials of evidence-based practice in education: A systematic review of randomised controlled trials in education research 1980-2016. *Educational Research*, 60(3), 276-291. <https://doi.org/10.1080/00131881.2018.1493353>
- Floretta, J., Friedlander, S., & Dhaliwal, I. (2020). Beyond Randomized Controlled Trials. *Stanford Social Innovation Review*. <https://doi.org/10.48558/XCWR-CW33>
- Freire, P. (1970). *Pedagogía del oprimido*. Siglo XXI Editores.
- Frontiers in Education. (2022). Evidence-based education: Objections and future directions. <https://www.frontiersin.org/articles/10.3389/feduc.2022.941410>
- Gorard, Stephen & See, Beng & Siddiqui, Nadia. (2017). The trials of evidence-based education: The Promises, Opportunities and Problems of Trials in Education. [10.4324/9781315456898](https://doi.org/10.4324/9781315456898)
- Hernández, R., & Mendoza, C. (2023). *Metodología de la investigación. Las rutas cuantitativa, cualitativa y mixta* (2da ed.). Ciudad de México: McGraw Hill. https://drive.usercontent.google.com/download?id=1JChN0wmF9WC57UOu_oRGXrahiy3Czw5V&export=download&authuser=0
- Hernández, R., Fernández, C., & Baptista, P. (2018). *Metodología de la investigación* (6ta ed.). México D.F, México, México: McGraw-Hill. <https://www.esup.edu.pe/wp-content/uploads/2020/12/2.%20Hernandez,%20Fernandez%20y%20Baptista-Metodolog%C3%ADa%20Investigacion%20Cientifica%206ta%20ed.pdf>
- J-PAL. (2024). Abdul Latif Jameel Poverty Action Lab. <https://www.povertyactionlab.org>
- Kaplan, A., Cromley, J., Perez, T., Dai, T., Mara, K., & Balsai, M. (2020). The role of context in educational RCT findings: A call to redefine “evidence-based practice”. *Educational Researcher*, 49(4), 285-288. <https://doi.org/10.3102/0013189X20921862>
- Kuhn, T. S. (1962). *The structure of scientific revolutions*. University of Chicago Press.
- Lortie-Forgues, H., & Inglis, M. (2019). Rigorous large-scale educational RCTs are often uninformative: Should we be concerned? *Educational Researcher*, 48(3), 158-166. <https://doi.org/10.3102/0013189X19832850>
- Magrath, B., Aslam, M., & Johnson, D. (2019). Systems research in education: Designs and methods. *Research in Comparative and International Education*, 14 (1). <https://doi.org/10.1177/1745499919828927>
- Martín Martínez, C. (2020). Efecto de intervenciones de actividad física en el ámbito escolar sobre las habilidades lingüísticas: una revisión sistemática y meta-análisis de ensayos controlados aleatorizados. Tesis de Grado. Universidad Complutense de Madrid. <https://hdl.handle.net/20.500.14352/8912>
- Nobel Prize. (2019). Understanding development and poverty: The experimental approach. <https://www.nobelprize.org/uploads/2019/10/advanced-economicsciencesprize2019.pdf>
- O'Connor, J. (2024). Randomised Controlled Trials in Education: An Experiment Worth Taking? *Irish Journal of Education / Iris Eireannach an Oideachais*, 48, 1-27. <https://www.jstor.org/stable/27332408>
- Odgaard-Jensen, J., Vist, G. E., Timmer, A., Kunz, R., Akl, E. A., Schünemann, H., Briel, M., Nordmann, A. J., Pregno, S., & Oxman, A. D. (2011). Randomisation to protect against selection bias in healthcare trials. *Cochrane Database of Systematic Reviews*. <https://doi.org/10.1002/14651858.MR000012.pub3>
- Odom, S. L. (2021). Education of students with disabilities, science, and randomized controlled trials. *Research and Practice for Persons with Severe Disabilities*, 46(3), 132-145. Odom, S. L. (2021). Education of Students with Disabilities, Science, and Randomized Controlled Trials. *Research and Practice for Persons with Severe Disabilities*, 46(3), 132-145. <https://doi.org/10.1177/15407969211032341>

- Olken, B.A. (2020). Banerjee, Duflo, Kremer, and the Rise of Modern Development Economics. *Scand. J. of Economics*, 122: 853-878. <https://doi.org/10.1111/sjoe.12418>
- Parra, J. D., & Edwards Jr, D. B. (2024). Challenging the gold standard consensus: Randomised controlled trials (RCTs) and their pitfalls in evidence-based education. *Critical Studies in Education*, 65(5), 513-530. <https://doi.org/10.1080/17508487.2024.2314118>
- Popper, K. (1959). The logic of scientific discovery. Hutchinson.
- Puente, J., Lucero, E., Díaz, L., Manuel, J., Ramírez, Q., & Román, D. (2024). Herramientas y metodologías para la evaluación científica: bibliometría, cienciométrica e informetría. *Seminars in Medical Writing and Education*, 3(395), 395. <https://doi.org/https://doi.org/10.56294/mw2024395>
- Ramírez, J. V., Robles, C. C., Macías, L. P., & Vega, D. M. (2020). Una unidad didáctica de Educación Deportiva como estrategia de educación para la ciudadanía en Educación Física: Un ensayo controlado aleatorizado por grupos. *Retos: nuevas tendencias en educación física, deporte y recreación*, (38), 44-52. <https://dialnet.unirioja.es/servlet/articulo?codigo=7397350&orden=0&info=link>
- Ramírez-Rivera, D. L., Martínez-Contreras, T., Ruelas, A. L., Henry-Mejía, G., Quizán-Plata, T., Esparza-Romero, J., ... & Díaz-Zavala, R. G. (2021). Efecto de una intervención en línea de cambio en el estilo de vida sobre el puntaje zIMC de escolares mexicanos: protocolo de ensayo controlado aleatorizado piloto cegado a evaluadores durante la pandemia por COVID-19. *Revista española de nutrición humana y dietética*, 25. <https://doi.org/10.14306/renhyd.25.S2.1320>
- Román, D. (2024). Terminología en metodología de la investigación cualitativa. In *Seminars in Medical Writing and Education*, 3(655), 655. <https://doi.org/https://doi.org/10.56294/mw2024655>
- Rowe, M., & Oltmann, C. (2016). Randomised controlled trials in educational research: Ontological and epistemological limitations. *African Journal of Health Professions Education*, 8(1), 6-8. Retrieved from <https://www.ajol.info/index.php/ajhpe/article/view/135389>
- Shalem, Y., & De Clercq, F. (2022). Randomized control trials in education (RCTs): What is their contribution to education theory about teaching?. *Journal of Education (University of KwaZulu-Natal)*, (89), 3-22. <https://doi.org/10.17159/2520-9868/i89a01>
- Slavin, R. E. (2020). How evidence-based reform will transform research and practice in education. *Educational Psychologist*, 55(1), 21-31. <https://doi.org/10.1080/00461520.2019.1611432>
- Smith, J., & Flanagan, S. (2024, 25 de septiembre). Repensar los ensayos controlados aleatorios. *Punto de Vista Medición y Evaluación SSIR*, (13). <https://ssires.tec.mx/es/noticia/repensar-los-ensayos-controlados-aleatorios>
- Spybrook, J., Zhang, Q., Kelcey, B., & Dong, N. (2020). Learning from cluster randomized trials in education: An assessment of the capacity of studies to determine what works, for whom, and under what conditions. *Educational Evaluation and Policy Analysis*, 42(3), 354-374. <https://doi.org/10.3102/0162373720929018>
- Staines, L., Healy, C., Corcoran, P., Keeley, H., Coughlan, H., McMahon, E., Cotter, P., Cotter, D., Kelleher, I., Wasserman, C., Brunner, R., Kaess, M., Sarchiapone, M., Hoven, C. W., Carli, V., Wasserman, D., & Cannon, M. (2023). Investigating the effectiveness of three school-based interventions for preventing psychotic experiences over a year period - a secondary data analysis study of a randomized controlled trial. *BMC Public Health*, 23, 219. <https://doi.org/10.1186/s12889-023-15107-x>
- Styles, B., & Torgerson, C. (2018). Randomised controlled trials (RCTs) in education research-methodological debates, questions, challenges. *Educational Research*, 60(3), 255-264.
- Styles, B., & Torgerson, C. (2018). Randomised controlled trials (RCTs) in education research - methodological debates, questions, challenges. *Educational Research*, 60(3), 255-264. <https://doi.org/10.1080/00131881.2018.1500194>
- The Conversation. (2019). 2019 Nobel Prize in Economics: The limits of the clinical trial method. <https://theconversation.com/2019-nobel-prize-in-economics-the-limits-of-the-clinical-trial-method-125888>
- Tyler, C., Akerlof, K. L., Allegra, A., Arnold, Z., Canino, H., Doornenbal, M. A., Goldstein, J. A., Pedersen, D. B., & Sutherland, W. J. (2023). AI tools as science policy advisers?

- The potential and the pitfalls. *Nature*, 622, 27-30. <https://doi.org/10.1038/d41586-023-02999-3>
- UNAM. (2004). Implicaciones de los paradigmas de investigación. https://www.revista.unam.mx/vol.5/num1/art1/ene_art1.pdf
- UNESCO. (2023). Global education monitoring report 2023 - Technology in education: A tool on whose terms? UNESCO. <https://www.unesco.org/gem-report/en/technology>
- Van Damme, D. (2022). The power of proofs: (Much) beyond RCTs. Centre for Curriculum Redesign. <https://curriculumredesign.org/wp-content/uploads/The-Power-of-Proofs.pdf>
- Welsh, R. O. (2021). Assessing the quality of education research through its relevance to practice: An integrative review of research-practice partnerships. *Review of Research in Education*, 45(1), 170-194. <https://doi.org/10.3102/0091732X20985082>
- What Works Clearinghouse. (2022). What Works Clearinghouse: Procedures and standards handbook, version 5.0. U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance (NCEE). <https://ies.ed.gov/ncee/wwc/Handbooks>
- Wolf, B., & Harbatkin, E. (2023). Making sense of effect sizes: Systematic differences in intervention effect sizes by outcome measure type. *Journal of Research on Educational Effectiveness*, 16(1), 134-161. <https://doi.org/10.1080/19345747.2022071364>
- World Bank. (2019). A Nobel Prize for development RCTs! <https://blogs.worldbank.org/en/impactevaluations/nobel-prize-development-rcts>

FINANCIAMIENTO

No aplica.

CONFLICTO DE INTERÉS

No aplica.

CONTRIBUCION AUTORAL:

Conceptualization: Jose Humberto Puente

Data curation: Jose Humberto Puente

Formal analysis: Jose Humberto Puente

Research: Jose Humberto Puente

Methodology: Jose Humberto Puente

Project management: Jose Humberto Puente

Resources: Jose Humberto Puente

Software: Jose Humberto Puente

Supervision: Jose Humberto Puente

Validation: Jose Humberto Puente

Display: Jose Humberto Puente

Drafting - original draft: Jose Humberto Puente

Writing - proofreading and editing: Jose Humberto Puente