

Revista de Estudios y Experiencias en Educación

REXE

journal homepage: <http://revistas.ucsc.cl/index.php/rexe>

Evaluaciones educativas a gran escala en los estados de Espírito Santo y Baja California: especificidades y aproximaciones entre Paebes y Exeims-BC

Ian Puppim Lopes, Ronildo Stieg, Wagner dos Santos
Universidade Federal do Espírito Santo, Vitória, Brasil


Recibido: 15 noviembre 2023 - Revisado: 21 de febrero 2024 - Aceptado: 14 de marzo 2024


RESUMEN


Este artículo tiene como objetivo analizar la trayectoria histórica de dos evaluaciones estandarizadas: el Programa de Evaluación de la Educación Básica de Espírito Santo (Paebes) en Brasil y el Examen Estatal de Ingreso a Educación Media Superior de Baja California (Exeims-BC) en México. Se caracteriza por ser una investigación cualitativa y descriptiva basada en el análisis documental y el método comparativo. Las fuentes estuvieron compuestas por Matrices de Referencia de evaluación elaboradas por las Secretarías de Educación de los respectivos estados. Los resultados muestran que en el estado de Espírito Santo, el Paebes ha sido adoptado como política de evaluación para el seguimiento de la educación básica desde el año 2000, desplegándose en diferentes aspectos para llegar a todo el alumnado de las etapas de educación primaria y secundaria. En cambio, el estado de Baja California oficializó el Exeims-BC en el año 2012, dirigido a todo el alumnado de secundaria. Se concluye que los impactos de Paebes han favorecido procesos de rendición de cuentas por encima de las políticas de bonificación de maestros, mientras que los resultados de Exeims-BC han generado rendición de cuentas a los estudiantes, generando una certificación utilizada para ingreso a la educación media superior del país.

Palabras clave: Evaluación estandarizada; rendición de cuentas; Brasil; México.

*Correspondencia: [Ronildo Stieg](mailto:Ronildo.Stieg@ufes.br) (R. Stieg).

 <https://orcid.org/0009-0008-6846-9186> (ian.lopes@edu.ufes.br).

 <https://orcid.org/0000-0001-8698-4087> (ronildo.stieg@ufes.br).

 <https://orcid.org/0000-0002-9216-7291> (wagner.santos@ufes.com).

Standardized educational assessments in the states of Espírito Santo and Baja California: specificities and similarities between Paebes and Exeims-BC

ABSTRACT

This article aims to problematize and understand the historical trajectory of two standardized evaluations: the Basic Education Evaluation Program of Espírito Santo (Paebes) in Brazil and the State Entrance Examination to Higher Secondary Education of Baja California (Exeims-BC) in Mexico. It is characterized by being a qualitative and descriptive research based on documentary analysis and the comparative method. The sources were composed of evaluation Reference Matrices prepared by the Secretariats of Education of the respective states. The results show that in the state of Espírito Santo, Paebes has been adopted as an evaluation policy for monitoring basic education since 2000, being deployed in different aspects to reach all students in the primary and secondary education stages. Conversely, the state of Baja California made the Exeims-BC official in 2012, aimed at all high school students. It can be concluded that the impacts of Paebes have promoted accountability processes related to teacher incentive policies, whereas the results of Exeims-BC have fostered student accountability, generating a certification used for entry into higher education in the country.

Keywords: Standardized assessment; accountability; Brazil; Mexico.

1. Introducción

Comprender los exámenes estandarizados como políticas educativas implica considerar que son instrumentos regulatorios impulsados por organismos gubernamentales y no gubernamentales, presente en diversos contextos (Álvarez-López y Matarranz, 2020; Monarca, 2020; Ozga, 2020; Verger et al., 2019). Aunque estas evaluaciones proporcionan datos valiosos para diagnosticar la situación de una realidad educativa específica, el uso de sus resultados para impulsar la mejora de la calidad y orientar las reformas educativas ha estado influenciado por el *Global Educational Reform Movement* (GERM). Según Carrasco y Gunter (2019), las ideas del GERM se han promulgado en diversos países, generando cambios en la oferta educativa, sobre todo impactando en cambios que incluyen la competencia entre escuelas, el enfoque en materias básicas y la rendición de cuentas basada en pruebas estandarizadas. Este enfoque gerencialista ha tenido un impacto considerable en la organización de los sistemas educativos, especialmente en países con sistemas escolares organizados en cuasi mercados, como en Chile (Verger et al., 2019).

Como señala Klein (2019), resulta necesario realizar análisis críticos para evaluar la efectividad de las metas establecidas para estas pruebas estandarizadas e identificar los ajustes necesarios. Por ende, en lugar de centrarse exclusivamente en la reproducción de objetivos, la generación de calificaciones y datos cuantitativos, las evaluaciones estandarizadas deben abarcar todos los elementos que componen el proceso educativo. Esto implica un enfoque más holístico que incluye la evaluación de sistemas, la evaluación institucional y la evaluación del proceso de enseñanza y aprendizaje. Este enfoque constituiría un proceso continuo e interconectado que fomente mejoras sustanciales en la calidad de la educación (Klein, 2019).

Mons (2009) sostiene que las evaluaciones educativas a gran escala son pruebas diseñadas para medir los logros cognitivos (conocimientos) del alumnado, cuyas características, aplicación, corrección tienden a ser uniformes. Estas pruebas consisten en monitorear el desempeño de estudiantes de una determinada edad/grado en una o más materias escolares. Sus resultados se analizan mediante comparaciones entre sistemas educativos para identificar posibles efectos de factores contextuales (escuela, estrategias pedagógicas y acción docente) sobre el rendimiento estudiantil en estas evaluaciones (Jiménez-Moreno, 2016; Mons, 2009).

Según Contreras y Torres (2023) y Ramos-Zincke (2018), el alineamiento de las evaluaciones estandarizadas con argumentos técnicos y científicos significa que estos exámenes se configuran como instrumentos de control en la definición de la calidad de las prácticas docentes, del aprendizaje y/o en la elaboración de perfiles de estudiantes/sociedad que se ajusten a las políticas educativas neoliberales actuales. Al asumir estas características, las evaluaciones estandarizadas, debido a los riesgos que conllevan sus objetivos, comienzan a adoptar lo que a nivel mundial se ha denominado sistemas de rendición de cuentas, cuyo enfoque se centra en medir el desempeño estudiantil en relación con ciertos contenidos curriculares (Vanlommel y Schildkamp, 2019).

Ozga (2020) define el término *accountability* como la responsabilidad impuesta a profesores y estudiantes de rendir cuentas de sus acciones ante un agente externo al entorno escolar. En consecuencia, la naturaleza de sus impactos puede presentarse de dos maneras: ofreciendo altos riesgos, con un impacto sustancial en la vida estudiantil, docente y/o en las instituciones escolares; o riesgos bajos, por lo tanto, no afectan directamente a los individuos involucrados, pero proporcionan información que respalda las políticas públicas en el área educativa (Marques et al., 2020).

Orfield y Wald (2000) explican que estos dos efectos están relacionados con la forma en que se aplican los resultados de estas evaluaciones, ya sea para promoción y certificación de estudiantes, o en bonos docentes como ocurre en diferentes estados brasileños cuyas políticas de evaluación han fomentado estas prácticas (Santos et al., 2021), o incluso en la acreditación o cierre de escuelas, como es el caso en contextos como Chile (Falabella y De La Vega, 2016).

Al analizar la producción académica que problematiza las evaluaciones estandarizadas y su papel político en los sistemas educativos, se evidencia que el énfasis de la investigación está: a) en los aspectos macrosociales discutiendo sus alcances (Falabella y De La Vega, 2016; Jiménez-Moreno, 2016; Smith, 2017); b) en el uso de este tipo de evaluación para satisfacer las demandas de una agenda política educativa global (Alkin y King, 2016; Beaver y Weinbaum, 2015; Brown y Zhang, 2017; Ebbeler et al., 2017; Sussman y Wilson, 2019; Vanlommel y Schildkamp, 2019); y c) en las percepciones/significados que directivos y administradores de redes educativas públicas y privadas atribuyen a estas evaluaciones (Contreras y Torres, 2023; Pino et al., 2016; Tolo et al., 2020). Ante este escenario, se observa un vacío en las investigaciones que analizan los exámenes estandarizados desde perspectivas microsociales (Ball, 2016), sobre todo, las evaluaciones estatales, que apuntan a comprender su surgimiento y organización.

En este sentido, el presente estudio plantea las siguientes interrogantes: ¿Qué políticas de evaluación han implementado los estados de Espírito Santo (Brasil) y Baja California (México) además de los exámenes estandarizados aplicados a nivel nacional e internacional? ¿Cuándo se pusieron en marcha y cuáles son las características de estas evaluaciones? ¿Qué similitudes y/o particularidades se pueden observar entre las evaluaciones desarrolladas en estos dos estados? Se plantea la hipótesis de que los estados de Espírito Santo y Baja California han adoptado políticas evaluativas estatales en respuesta a la búsqueda de mejoras en

el rendimiento académico de los estudiantes en exámenes estandarizados a nivel nacional e internacional, alineándose con la tendencia observada en otros estados (Bauer et al., 2015; Pontes Júnior et al., 2021).

Partiendo del supuesto de que estos exámenes tienen impactos (Afonso, 2009; Orfield y Wald, 2000) en el escenario educativo de los estados de Espírito Santo y Baja California, respectivamente, este artículo tiene como objetivo problematizar y comprender la trayectoria histórica de las políticas de evaluación implementadas en estos dos estados. Específicamente, analiza las características de cada examen y su estructura organizacional; las materias evaluadas y los formatos de las evaluaciones.

2. Metodología

2.1. Enfoque

Se caracteriza por ser una investigación documental de naturaleza cualitativa (Creswell y Clark, 2013) y se fundamenta en el método comparado en educación (Caballero et al., 2016; García Garrido, 1996). Según García Garrido (1996), este método consta de al menos cinco fases: I) *Identificación del problema y formulación de hipótesis* (corresponde a la presentación de la pregunta educativa y la expectativa de solución para dicho problema); II) *Delimitación de la investigación* (realidad educativa investigada para la comparación, pudiendo ser a nivel local, nacional o internacional o a lo largo del tiempo); III) *Estudio descriptivo-interpretativo o analítico* (corresponde a la descripción de la información recolectada, su clasificación y las estrategias de presentación elegidas, e interpretaciones de las causas y/o factores y/o características comunes que se relacionan con el objeto de este estudio); IV) *Yuxtaposición y comparación* (presentación de los datos en paralelo y haciendo comparaciones para dar a conocer las semejanzas y diferencias); V) *Prospectiva* (conclusiones del estudio y proposición de algunas líneas de acción).

Según Ferrer (2002), los estudios enfocados en la educación comparada forman parte de un plan más amplio para explicar los fenómenos presentes en los sistemas educativos y su vinculación con el ambiente social. Por ello, el estudio de sociedades distintas como México y Brasil, que comparten el fenómeno de las políticas de evaluación a gran escala a nivel estatal, se convierte en un objeto de estudio interesante para contribuir a investigaciones con un enfoque de educación comparada (Álvarez-López y Matarranz, 2020; Caruso, 2011; García Garrido et al., 2005; Navarrete-Cazales y Rojas Moreno, 2022).

2.2. Unidades de análisis

Inicialmente se partió de la base de datos existente, construida en la Universidad Federal de Espírito Santo (UFES) y la Universidad Autónoma de Baja California (UABC) a través del proyecto con financiación FAPES n° 11367/2021. Se utilizaron como fuentes las Guías Informativas y Matrices de Referencia de las evaluaciones de los estados de Espírito Santo (ubicado en la región sureste de Brasil) y Baja California (ubicado en el noroeste de México), producidas por las Secretarías de Educación de los respectivos estados, con el objetivo de regular y orientar el Programa de Evaluación de la Educación Básica de Espírito Santo (Paebes) y el Examen Estatal de Ingreso a Educación Media Superior de Baja California (Exeims-BC).

El relevamiento de fuentes se realizó en los sitios web de las Secretarías de Educación del Estado de Espírito Santo y Baja California. Además, se realizaron reuniones presenciales con Directivos de la Secretaría Estatal de Educación de Espírito Santo (SEDU) en Brasil para preguntar sobre el sistema de organización de Paebes, y con el personal docente de investigación de la UABC en México, a través de videoconferencias, para recopilar información adicional sobre Exeims-BC. Todo este proceso se desarrolló de marzo a agosto de 2022.

A partir de las fuentes documentales y la información proporcionada en las reuniones, se logró crear una base de datos con información que ofrecía pistas (Ginzburg, 1989) sobre cuándo surgieron Paebes y Exeims-BC, cómo se han constituido desde sus inicios y la forma en que se estructuran, incluyendo el público objetivo, las materias evaluadas y el formato de las pruebas.

2.3. Técnicas de recolección de datos

En las búsquedas en los sitios web de las Secretarías de Educación de los respectivos estados (Espírito Santo y Baja California), se emplearon diversos descriptores para encontrar las nomenclaturas correspondientes, tal como se detalla en la Tabla 1.

Tabla 1

Descriptores utilizados en los sitios web.

Brasil	México
<i>Sistema de avaliação em nível estadual</i>	Sistema de evaluación a nivel estatal
<i>Avaliação estadual para ensino médio</i>	Evaluación estatal para la enseñanza secundaria
<i>Teste educacional padronizado para ensino médio</i>	Prueba educativa estandarizada para la enseñanza secundaria
<i>Teste estadual no Espírito Santo</i>	Prueba de estado de Baja California
<i>Exame estadual para os alunos do ensino médio</i>	Examen estatal para estudiantes de secundaria

Fuente: Elaboración propia, 2023.

Los descriptores se utilizaron en función del idioma oficial de los países (portugués en Brasil y español en México). Además, es importante destacar que al analizar los exámenes de países como Brasil y México, se pueden observar algunas particularidades en relación a la estructura organizacional de los sistemas educativos de educación básica, así como el uso de diferentes terminologías debido al contexto sociocultural en el que se desarrollan. En el análisis, se consideran sistemas específicos (Tabla 2). En este sentido, se buscó analizar los exámenes estatales sin desconocer las delimitaciones conceptuales asumidas en cada país.

Tabla 2

Estructura de la educación básica obligatoria en Brasil y México.

BRASIL	MÉXICO
Etapas de enseñanza	Nivel educacional
Educación infantil (Preescolar) Duración de 3 años	Inicial o Preescolar Duración de 3 años
Educación elemental I y II Duración de 9 años	Primario Duración de 6 años
Escuela secundaria Duración de 3 años	Secundario Duración de 4 años
Educación Profesional Técnica Secundaria Duración de 4 años	Educación Secundaria Superior Duración de 3 años

Fuente: Elaboración propia, basado en información de los Ministerios Nacionales de Educación de [Brasil](#) y [México](#) (2023).

Según la Tabla 2, se observa que los sistemas de educación básica de Brasil y México presentan similitudes al analizar las tres primeras etapas/niveles de la educación, las cuales son obligatorias, pero difieren en la duración de algunas de ellas. En ambos sistemas educativos, la educación primaria es la etapa de mayor duración. Además, los últimos años de las escuelas en ambos países tienen una duración similar, pero con estudiantes de diferentes edades: en Brasil, el alumnado tiene entre 15 y 17 años, y en México, entre 12 y 16 años. Para este estudio, se consideraron estas estructuras de educación básica para analizar las evaluaciones estandarizadas que se realizan en al menos dos etapas/niveles de educación: primaria y secundaria.

2.4. Procesamiento de análisis

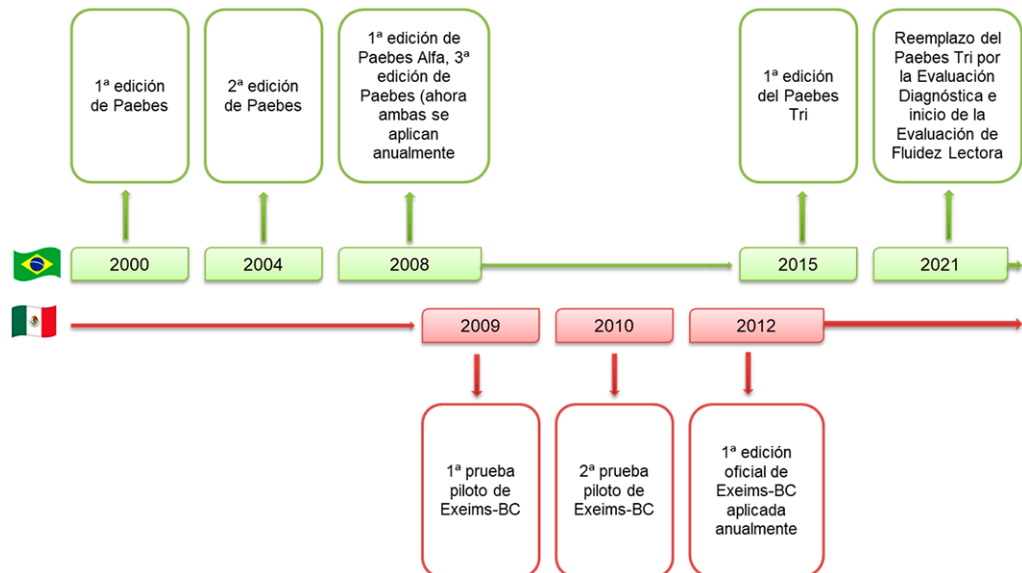
Para analizar las fuentes, se empleó el software Microsoft PowerPoint para generar la Figura 1, que representa una línea de tiempo de Paebes y Exeims-BC. Esta figura considera los años de implementación, las continuidades y discontinuidades de las políticas de evaluación realizadas en los estados de Espírito Santo y Baja California en relación a sus fases y desarrollos a lo largo de los años. Asimismo, se produjo la Tabla 3 para discutir las materias evaluadas y el formato de sus últimas ediciones. Se definieron los siguientes ítems de análisis: a) la trayectoria de los exámenes; b) las características generales de cada examen; y c) la comparación de evaluaciones estatales para abordar las especificidades y similitudes entre ellas.

3. Resultados

El análisis de los documentos relacionados con Paebes y Exeims-BC permitió la creación de la Figura 1. Se identificó que, particularmente en el estado de Espírito Santo, el Paebes ha sido adoptado como política de evaluación educativa desde el año 2000. Del mismo modo, en el estado de Baja California, se implementó el Exeims-BC como política de evaluación, y desde 2012, este examen ha generado sistemas de certificación utilizados para la selección de estudiantes que ingresarán a cursos de nivel medio superior (bachillerato) en todo el país.

Figura 1

Trayectoria de los exámenes Paebes y Exeims-BC.



Fuente: Elaboración propia, basado en información de las Secretarías de Educación de [Espírito Santo](#) y [de Baja California](#) (2023).

Con base en la Figura 1, se observa la evolución y expansión del Paebes en Espírito Santo desde su establecimiento en 2000, con el propósito fundamental de evaluar el rendimiento de la red pública estatal en la educación primaria y secundaria (Espírito Santo, 2022). Se destaca que, aunque inicialmente se realizó de manera intermitente entre 2000 y 2008, se convirtió en el primer examen en tener una edición oficial al compararlo con el Exeims-BC. A partir de la segunda edición en 2004, el Paebes incorporó la metodología estadística conocida como Teoría de Respuesta al Ítem, que busca calcular la probabilidad de que los estudiantes respondan correctamente a un ítem específico. Esta metodología ha sido empleada consistentemente en las evaluaciones más recientes (Marques et al., 2020). La transición a una frecuencia anual del Paebes desde 2008 se originó a partir de diversas iniciativas implementadas por el gobierno estatal, incluyendo el Plan Estratégico de la Nueva Escuela (Espírito Santo, 2008), la introducción del nuevo currículo básico común basado en competencias y habilidades (Ordenanza n° 143-R, de 17 de noviembre de 2009), y el establecimiento del Índice de Desarrollo de la Educación Básica en Espírito Santo (IDEB) (Ley Complementaria n° 504, de 23 de noviembre de 2009). Estas decisiones resultaron en una mayor inversión en el sector educativo por parte del estado de Espírito Santo, especialmente mediante la contratación del *Centro de Políticas Públicas e Avaliação da Educação* (CAEd) de la Universidad Federal de Juiz de Fora para administrar el Paebes y otras evaluaciones relacionadas (Côco y Gontijo, 2017). Cabe destacar la implementación de nuevas evaluaciones, como el Paebes Alfa en 2008, el Paebes Tri en 2015, junto con las evaluaciones Diagnóstica y de Fluidez Lectora en 2021.

Se destaca que el Paebes ha tenido una influencia significativa en las políticas educativas de Espírito Santo, debido al alcance, profundidad y regularidad de las evaluaciones estatales. Para organizar y producir estos exámenes, el estado estableció una alianza con el CAEd en 2015, una entidad que apoya la estructuración e implementación de evaluaciones estandarizadas en distintos ámbitos públicos, incluyendo los departamentos de educación municipales y estatales. En 2021, mediante el Decreto n° 4889-R, se creó la Gerencia de Evaluación (GEA) con la función de planificar, producir, almacenar y proporcionar información sobre evaluaciones educativas en Espírito Santo, a nivel estatal, nacional e internacional (Espírito Santo, 2022).

Con la asistencia del CAEd en la operacionalización del Paebes, los resultados de las evaluaciones estatales en Espírito Santo, aunque restringidos al CAEd y a los directores y docentes de las escuelas, y no públicos a toda la comunidad civil, sirvieron como subsidio para la (re)formulación de políticas educativas con el objetivo de mejorar la calidad de la educación en el estado (Espírito Santo, 2022).

Se observa, según la Figura 1, que en 2009 se produjo un punto de encuentro entre las evaluaciones estandarizadas en Espírito Santo y Baja California. En ese año, Paebes Alfa comenzó a abarcar un mayor número de municipios en el estado y se realizó la primera prueba piloto del Exeims-BC. Es notable que, aunque el Paebes no fue considerado una prueba piloto en Espírito Santo, a diferencia del examen de Baja California, sus primeras ediciones no tuvieron cobertura universal en el estado, aunque con el tiempo obtuvieron aceptación en los municipios (Côco y Gontijo, 2017).

En el caso del Exeims-BC, se llevaron a cabo dos pruebas con grupos de muestra de estudiantes en los años 2009-2010. Posteriormente, se dedicó un periodo para analizar y mejorar el examen en 2011, lo que permitió la realización de la primera edición oficial en 2012. Esta edición tenía el propósito abarcar a estudiantes de secundaria en todos los municipios del estado de Baja California. A partir de ese año, el examen se lleva a cabo anualmente (Baja California, 2012; México, 2022). Desde entonces, los datos generados por el Exeims-BC se han utilizado como un mecanismo de selección de estudiantes al finalizar la educación secundaria, con el objetivo de ingresar a la educación secundaria superior. Este examen, de

naturaleza mixta y estandarizada, tiene un carácter diagnóstico para selección y nivelación. Es ofrecido y preparado por el Sistema Educativo del Estado (SEE), a través de la Dirección de Evaluación Educativa del estado de Baja California. En cuanto al objetivo del Exeims-BC, según la información disponible en el SEE, busca "(...) proporcionar y promover los servicios de educación, cultura y deporte con equidad, calidad y sentido humano, en la formación para la vida de todos los habitantes de Baja California" (México, 2022, p. 1).

Al comparar los años de implementación, continuidades y discontinuidades, se destacan las principales diferencias entre los dos exámenes (Figura 1): a) ambos fueron implementados en años diferentes, el Paebes en 2000 y el Exeims-BC en 2012; b) Paebes se realizó cada cuatro años en sus tres primeras ediciones, mientras que el Exeims-BC se ha llevado a cabo anualmente desde su creación; c) antes de la primera edición oficial del Exeims-BC, se realizaron dos pruebas piloto, lo que no ocurrió con el Paebes; d) desde 2000 hasta 2008, solo existió el Paebes (realizado a finales de año), pero en 2008 se inició Paebes-Alfa, seguido de Paebes Tri en 2015 y Evaluaciones Diagnósticas y Fluidez Lectora en 2021. Es importante resaltar que en 2021 se puso a disposición, por primera vez, el Cuestionario Socioemocional como un aspecto evaluado dentro de la Evaluación Diagnóstica en Espíritu Santo.

Respecto a las similitudes entre el Paebes (con sus distintos aspectos) y el Exeims-BC, se identifica que ambos: a) se llevan a cabo anualmente a partir de un período determinado (el Paebes desde 2008 y el Exeims-BC desde 2012); b) tomaron la demanda internacional como uno de los puntos de partida para la creación de un programa estatal de evaluación estandarizada; c) evalúan a estudiantes de educación básica; d) se desarrollan en países latinoamericanos con características de estados federados; e) no se realizaron en 2020 debido a la Covid-19.

Es importante comprender las motivaciones que guían las decisiones de los respectivos gobiernos estatales a lo largo de los años, ya que ambos exámenes comenzaron a administrarse anualmente, sin seguir el modelo escalonado de las primeras pruebas de Paebes. Se destaca que este tipo de cambios funciona como un mecanismo de control estatal para evaluar la calidad de la educación en la red estatal, obtener datos con mayor frecuencia y, en consecuencia, impulsar reformas en las políticas educativas (Cunha y Müller, 2018).

En el caso de Paebes, la secuencia de pruebas previas y su posible éxito en su objetivo en las escuelas de la muestra parecen haber sido decisivos para desencadenar los cambios y la ampliación del programa (Còco y Gontijo, 2017). Sin embargo, Machado et al. (2015) señalan que un factor impulsor en el desarrollo de exámenes estandarizados estatales en los países latinoamericanos, especialmente en Brasil, nació bajo la influencia de las políticas de evaluación externa promovidas por el Ministerio de Educación (MEC), con una notable participación de organismos internacionales, entre ellos el Banco Mundial. Por otro lado, la aparición tardía de una evaluación estatal en Baja California (2012) en México ofrece pistas de un movimiento de resistencia a los exámenes estandarizados en ese país.

En la Figura 1 se identifican las diferentes versiones del Paebes. Por ejemplo, el Paebes Alfa ha funcionado como una evaluación diagnóstica del proceso de alfabetización en los primeros años de educación primaria (1° a 3°). Por otro lado, el Paebes Tri se caracterizó por ser una evaluación continua desde su creación hasta su finalización (2015-2019); solía realizarse una vez por trimestre académico y sus resultados se utilizaban para planificar estrategias con el fin de mejorar el rendimiento del alumnado en ediciones posteriores. Esta evaluación era aplicada y corregida por el personal docente de las propias unidades escolares. Además, la calificación obtenida en el Paebes Tri se sumaba a las evaluaciones de aprendizaje realizadas por el personal docente de las materias evaluadas en el contexto escolar al final de cada trimestre (Santos et al., 2021).

En 2021, Espíritu Santo lanzó la Evaluación de Fluidez Lectora, que se lleva a cabo anualmente al final del segundo trimestre académico. Esta evaluación se realiza “(...) a través de una aplicación propia, adecuada para Smartphone o tabletas, capaz de registrar la lectura de palabras, pseudopalabras y textos narrativos realizada por el alumnado (Espíritu Santo, 2023, p. 10, nuestra traducción).

Ese mismo año, también se lanzó la Evaluación Diagnóstica, en reemplazo del Paebs Tri y se lleva a cabo al inicio del año escolar y al final del segundo trimestre. Su objetivo principal es diagnosticar los aprendizajes y habilidades que el alumnado ha desarrollado y aquellos que aún no ha consolidado. Esto ayuda en la planificación de los equipos pedagógicos escolares, las Superintendencias Regionales de Educación (SRE) y la gestión de la SEDU. Además, sirve como base para las propuestas de intervención pedagógica en el año escolar. Estos cambios en el Paebs fueron anunciados en las ordenanzas n° 138-R y n° 139-R, publicadas en junio de 2022 (Espíritu Santo, 2022).

Además de la trayectoria de Paebs y Exeims-BC, otro punto fundamental para comprender estos exámenes es la estructura de las pruebas, que tienen diferentes configuraciones según su finalidad. Se observan diferencias en las etapas/niveles de enseñanza y en las materias evaluadas, años de aplicación y formato de sus últimas ediciones (Tabla 3). Parte superior del formulario

Tabla 3

Características generales de las evaluaciones Paebs y Exeims-BC.

	Espíritu Santo (ES)				Baja California (BC)	
	Paebs	Paebs Alfa	Paebs Tri	Evaluación Diagnóstica	Evaluación de la Fluidez Lectora	Exeims-BC
Etapa/ nivel de educación evaluado	5° y 9° año de primaria y 3° de secundaria	1°, 2° y 3° de primaria	1°, 2°, 3° año de secundaria y 4° año de secundaria integrado	3° a 9° año de primaria y 1° a 3° año de secundaria	2° de primaria	1°, 2° y 3° de Secundaria
Años de aplicación	De 2000 a 2019 En 2021 y 2022	De 2008 a 2019 En 2021 y 2022	De 2015 a 2019	En 2021 y 2022		De 2012 a 2019 En 2021 y 2022
Materias evaluadas	Lengua Portuguesa, Matemáticas, Biología, Historia y Geografía	Lengua Portuguesa y Matemáticas		Lengua Portuguesa, Matemáticas, Historia, Geografía, Biología, Física y Química	Lengua Portuguesa	Matemáticas, Lengua Española, Inglés, Física, Química, Biología, Humanidades, Ciencias Sociales, Cívica y Ética
Formato		Impreso		Digital e impreso		Digital

Fuente: Elaboración propia, basado en información de las [Secretarías de Educación de Espíritu Santo y de Baja California \(2023\)](#).

En cuanto a las etapas/niveles educativos evaluados, se observa en la Tabla 3 que cada evaluación tiene su público específico. El Paebes se centra en evaluar la adquisición y consolidación de competencias al finalizar las etapas de 5º y 9º de primaria y 3º de secundaria. Por otro lado, el Paebes Alfa está dirigido a alumnos de 1º a 3º de primaria y se enfoca en evaluar la alfabetización, similar a la Evaluación de Fluidez Lectora que realiza el alumnado de 2º de primaria. Las diferencias entre estos exámenes radican en el formato y las materias evaluadas. El Paebes Alfa es en versión impresa, mientras que la Evaluación de Fluidez Lectora es digital. Además, el primero evalúa conocimientos de lengua portuguesa y matemáticas, mientras que el segundo se centra exclusivamente en lengua portuguesa.

En el caso de la Evaluación Diagnóstica, su alcance se extiende a los tres grados de secundaria regular y el 4º de secundaria integrado. De todas las evaluaciones estandarizadas en el sistema de educación básica de Espírito Santo, dos de ellas tienen un impacto más directo en la vida de los profesionales de la educación (Marques et al., 2020; Vanlommel y Schildkamp, 2019). Tanto el Paebes como el Paebes Alfa son la base para calcular: a) el Bono de Desempeño para los profesionales de SEDU; b) el Premio Escuela que Colabora, que otorga premios económicos a 50 escuelas públicas con los mejores promedios de Paebes y apoya a 50 escuelas con calificaciones más bajas según el IDEB (Espírito Santo, 2023; Soares et al., 2022).

Respecto al Exeims-BC, se observan, con base en la Tabla 3, características similares a la Evaluación Diagnóstica en cuanto a: a) el nivel educativo evaluado (con un enfoque mayor en educación secundaria); b) un conjunto más amplio de materias evaluadas; y c) el formato de las pruebas (digital). En general, es posible comprender que las evaluaciones en Espírito Santo son más integrales y diversificadas en cuanto a las etapas/niveles de enseñanza y las materias evaluadas. Además, los formatos son variados en comparación con el Exeims-BC de Baja California (COBACHBC, 2020).

También es notable el caso de la Evaluación Diagnóstica, iniciada en 2021, que, además de evaluar componentes curriculares de la formación general básica, incluye un Cuestionario Socioemocional, que funciona como una autoevaluación orientada a proporcionar información sobre "(...) competencias socioemocionales con el fin de intervenir psicopedagógicamente en el trabajo escolar con el alumnado (Espírito Santo, 2023, p. 8, nuestra traducción).

Si, por un lado, la continuidad de los exámenes en los dos contextos analizados presentó dinámicas diferentes, ya sea manteniendo una política de evaluación estatal única como el Exeims-BC o con ampliación de estas evaluaciones con formas nuevas y cada vez más integrales como el Paebes, también es evidente que, por otra parte, en Baja California y Espírito Santo, los conocimientos/materias predominantes en las evaluaciones estatales siguen siendo lengua materna (portugués o español) y las matemáticas.

En este contexto, se coincide con Menegão (2016) en la importancia de reconocer que los sistemas de evaluación, incluyendo Paebes y Exeims-BC, imponen un control curricular que favorece un conjunto específico de conocimientos, limitando la perspectiva educativa. Expertos en evaluación han señalado esto durante más de dos décadas (Bauer et al., 2013; Gentili y Silva, 1994), lo que expresa la naturaleza neoliberal de estas políticas (Álvarez-López y Matarranz, 2020; Monarca, 2020; Ozga, 2020; Verger et al., 2019), las cuales son conformadas por el nuevo orden internacional, que determina qué materias son prioritarias para la evaluación. A pesar de que se mencionan otras disciplinas en las fuentes, es sorprendente que hasta el año 2021, materias como Sociología, Filosofía, Artes y Educación Física aún no se hayan incorporado a estas evaluaciones.

4. Discusión

El análisis de dos políticas de evaluación estatales (Espíritu Santo y Baja California) evidencia características donde hay inclinación de Paebes hacia el Movimiento Mundial de Reforma Educativa (GERM), ya que parece servir a un proceso de rendición de cuentas con descentralización de la autonomía escolar, al transferirla al sector privado a través de un contrato financiado con dinero público, preparando y generando resultados a partir de evaluaciones realizadas a nivel estatal o nacional (Carrasco y Gunter, 2019; Jiménez-Moreno, 2016; Verger et al., 2019). Este proceso ha sido entendido por Carrasco y Gunter (2019) como principios políticos del GERM, ya que la provisión gubernamental de funciones y servicios se traspasa, en el caso del estado de Espíritu Santo, a CAEd, caracterizado según Verger, et al. (2017) como privatización y/o venta de servicios públicos al sector privado.

Aunque el Exeims-BC es una evaluación estatal reciente (2012), su principal objetivo es obtener un diagnóstico de la educación y, sobre todo, promover la certificación (Orfield y Wald, 2000) para que el alumnado de secundaria pueda acceder a la etapa educativa media superior, de dos años de duración (correspondiente al curso técnico en Brasil). En este caso, para seguir una carrera universitaria, el alumnado deberá completar todas las etapas educativas básicas, incluida la educación media superior.

La implementación del Exeims-BC realiza inferencias cuantitativas y cualitativas, que retroalimentan la toma de decisiones de las autoridades educativas del subsistema al que pertenece, además de brindar al personal docente y directivos información para trazar estrategias situadas que promuevan el fortalecimiento y consolidación de las competencias educativas establecidas para el alumnado (Jiménez-Moreno, 2016; Smith, 2017). En consecuencia, el Exeims-BC contribuye al sistema educativo, caracterizando el nivel del alumnado de secundaria en Baja California y fortaleciendo el bachillerato. El formato digital del examen ha permitido a profesores, familias y personas investigadoras consultar los resultados, manteniendo la confidencialidad de los implicados (Baja California, 2012). Por otro lado, faltan documentos normativos y estudios relacionados con el examen que ayuden a una comprensión más amplia de sus implicaciones en los procesos de rendición de cuentas (Loeb y Byun, 2019; Ozga, 2020).

A diferencia del Exeims-BC, el Paebes ha consolidado su posición como una evaluación educativa estandarizada en el estado de Espíritu Santo a lo largo de sus 12 ediciones. Su objetivo principal es evaluar el rendimiento del alumnado en la educación primaria y secundaria de la red pública estatal (Espíritu Santo, 2023). A pesar de algunas preocupaciones sobre la transparencia de información, SEDU ha utilizado los resultados del Paebes para identificar tanto las debilidades como las fortalezas de las instituciones educativas (Santos et al., 2021). Estos datos han sido fundamentales "(...) para gestores y docentes en la planificación de acciones pedagógicas orientadas a mejorar el aprendizaje, y proporcionan indicadores que orientan la implementación, (re)formulación y seguimiento de políticas educativas orientadas a promover la equidad y la calidad de la educación en Espíritu Santo" (Espíritu Santo, 2023, p. 13, nuestra traducción).

Sobre este tema, el gobierno federal promulgó la Ley n° 12.527/2011, conocida como Ley de Acceso a la Información (LAI) (Brasil, 2011), que reguló el derecho constitucional de acceso a la información pública. Esta legislación introdujo mecanismos que permiten a cualquier ciudadano obtener información pública de órganos y entidades gubernamentales sin necesidad de justificar la solicitud. La LAI establece que los organismos públicos deben divulgar información de interés público y carácter colectivo, incluso sin solicitudes específicas, respondiendo a preguntas comunes de la sociedad. Sin embargo, al analizar los sitios web de la SEDU y del Paebes, se encontraron páginas deficientes en información, algunas con

contenido estático y superficial sobre Paebes. Esto ha generado una alta demanda de transparencia, donde los ciudadanos deben solicitar información al CAEd, en lugar de beneficiarse de la transparencia activa que brindaría acceso fácil a la información en los sitios web oficiales de las entidades públicas.

Tras analizar los años desde la creación e implementación de Paebes y Exeims-BC, es evidente que la política de evaluación en Espírito Santo ya se estaba fortaleciendo durante un período de 12 años antes de que se estableciera el primer examen estatal en Baja California. Estos datos también indican que en Brasil, las políticas de evaluación se han fortalecido a nivel nacional desde la década de 1990, marcándose especialmente con las primeras evaluaciones como el Examen Nacional de Enseñanza Media (Enem) en 1998 (Cunha y Müller, 2018), seguido de la primera edición de Paebes dos años después. No obstante, como en otros países latinoamericanos, Brasil parece seguir un modelo de política educativa influenciado por el Programa de Reforma Educativa en América Latina y el Caribe (Preal), financiado por el Banco Mundial. El enfoque principal de este programa es implementar y supervisar una agenda política para el desarrollo de la educación básica en la región (Gajardo, 1999).

En el caso de Brasil, según Míguez (2023), las reformas educativas fueron impulsadas por múltiples organismos internacionales que buscaban fortalecer las nuevas democracias emergentes. Entre las justificaciones estaba la necesidad de incrementar la calidad de la educación y universalizar el acceso a los sectores más desfavorecidos. Como consecuencia de este proceso, surgieron evaluaciones estandarizadas para monitorear y promover un conjunto de sistemas educativos (Míguez, 2023).

Como resultado, se observa una proximidad al comparar las fechas de implementación de Paebes y Exeims-BC, ya que ambos siguen el modelo de políticas nacionales de evaluación. Paebes se asemeja al Enem y Exeims-BC establece similitudes con la Evaluación Nacional del Logro Académico en Centros Escolares (Enlace), creada en 2006.

En un contexto dominado por políticas de evaluación, se aprecia que, particularmente en Espírito Santo, el alumnado participa anualmente en diversas evaluaciones, destacadas en este estudio a nivel estatal (Evaluaciones Diagnóstica, Evaluación de Fluidez Lectora, Paebes y Paebes Alfa), así como a nivel nacional (Saeb, Enem y Encceja) e internacional (Erce, Pisa, PIRLS y TIMSS) (Espírito Santo, 2023). La justificación para el excesivo número de exámenes promovidos por SEDU es que "(...) son necesarios para monitorear el funcionamiento de las redes educativas y brindar apoyo a sus gestores en la formulación de políticas educativas basadas en evidencia (Espírito Santo, 2023, p. 4, nuestra traducción).

La introducción de evaluaciones adicionales a nivel estatal en Espírito Santo y la contratación de organizaciones especializadas, como CAEd, responsable por producir, supervisar y utilizar un volumen cada vez mayor de datos recopilados, ha producido el incremento de un sistema de rendición de cuentas que muchas veces implica una compleja red de relaciones, interdependencias y reciprocidades entre el Estado evaluador y agentes externos en la formulación de evaluaciones a gran escala (Ozga, 2020). En consecuencia, el impacto de esto en la vida de los docentes ha generado una dimensión de responsabilización particularmente delicada, ya que tienen que rendir cuentas ante diversos órganos jerárquicos, desde el ministerio hasta las secretarías, los directores de escuela, los padres y los propios estudiantes (Afonso, 2009; De la Vega y Picazo, 2016). Verger, Fontdevila y Parcerisa (2019) señalan que esta dinámica se asemeja a la estructura de las evaluaciones a gran escala presentes en otros contextos, sobre todo en territorio inglés.

Se comprende, según Afonso (2009), que los exámenes estandarizados, aunque están diseñados y basados en la perspectiva de la responsabilidad y la rendición de cuentas, a menudo se centran en atender demandas externas a la escuela, pasando por alto el progreso académico y los procesos de aprendizaje del alumnado.

En este sentido, se observa que la centralidad otorgada por los exámenes estandarizados, principalmente en Paebes, sigue una lógica interna que se difunde ampliamente en las evaluaciones educativas estandarizadas de otros países. Se les atribuye un papel regulador, especialmente cuando se convierten en dispositivos con altas consecuencias, fundamentados en el modelo de rendición de cuentas (Contreras y Torres, 2023; Loeb y Byun, 2019; Ramos-Zinke, 2018; Suspitsyna, 2010). Como consecuencia, los resultados de estos exámenes suelen estar asociados con: a) la certificación del alumnado (Orfield y Wald, 2000) y la definición de a qué carreras podrán acceder en el nivel secundario superior, como se observa en México; b) incentivos y sanciones para docentes e instituciones en forma de beneficios económicos, como ocurre en el estado de Espíritu Santo (Santos et al., 2021); c) el cierre de escuelas, como sucede en Chile (De La Vega y Picazo, 2016; Elacqua et al., 2016; Falabella y De La Vega, 2016).

Aunque el estado de Baja California presenta un considerable retraso en responder a las políticas internacionales relacionadas con evaluaciones estandarizadas, el Espíritu Santo, por el contrario, ha buscado satisfacer esta demanda en un corto espacio de tiempo mediante diversas iniciativas, incluyendo la creación de más evaluaciones dentro de la agenda escolar. Hay indicios de que esta acción se justifica debido a la inversión financiera que se incrementaría según los resultados de cada país de América Latina, especialmente en base a las métricas establecidas por las evaluaciones internacionales. Sin embargo, es necesario considerar los riesgos que estas estrategias de evaluación mediante la rendición de cuentas pueden implicar. Según Verger et al. (2019), en países como Chile, se han adoptado evaluaciones nacionales a gran escala y medidas de rendición de cuentas en combinación con políticas de privatización exógenas (como *vouchers* u otras formas de subsidios públicos para escuelas privadas), con el objetivo de estimular la competición en educación.

No obstante, es necesario preguntarse hasta qué punto la concepción subyacente a este exceso de evaluaciones estandarizadas está al servicio de un currículo instrumental, cuyo modelo educativo está orientado a países en desarrollo, influenciado por organismos internacionales como el Banco Mundial, la OCDE y la UNESCO. En las últimas décadas, la OCDE y los países de ingresos medios han otorgado una mayor importancia a las evaluaciones a gran escala, los objetivos, los estándares y la rendición de cuentas en la gobernanza de sus sistemas educativos (Jiménez-Moreno, 2016; Ozga, 2020; Verger et al., 2019). Según Libâneo y Freitas (2018), este diseño curricular prioriza la preparación inmediata para el mercado laboral, con énfasis en habilidades prácticas y evaluaciones basadas en pruebas, ya sea a nivel local (estatal), nacional o internacional, con el principal objetivo de responsabilizar a la escuela y al personal docente por el desempeño del alumnado.

En este sentido, los países con mejores resultados educativos pueden haber ejercido presión sobre algunos países latinoamericanos, incluido Brasil, para adoptar estrategias similares mediante la implementación de nuevos exámenes con el fin de enfatizar una cultura de preparación para evaluaciones externas internacionales. Al mismo tiempo, otros países, como México, por ejemplo, pueden manifestar resistencia a estas políticas, postergando la implementación de nuevas evaluaciones, principalmente estatales.

5. Conclusiones

Del análisis de las Guías Informativas y Matrices de Referencia del Paebes y el Exeims-BC, se destaca que ambos son exámenes estandarizados estatales en países federativos. No obstante, difieren en objetivos, alcance y duración. El Paebes evalúa el desempeño en educación primaria y media, mientras que el Exeims-BC, como examen mixto, certifica el perfil de finalización (Orfield y Wald, 2000) en secundaria, brindando indicadores para el ingreso a la educación media superior en Baja California. Aunque ambos aportan a las políticas educativas, el Exeims-BC no lo hace como enfoque principal, a diferencia del Paebes.

El Paebes existe por más de dos décadas, contrastando con los poco más de 10 años del Exeims-BC. El progreso y expansión son notables en el Paebes, que ha evolucionado con el Paebes-Alfa y otras evaluaciones, a diferencia del Exeims-BC. Sin embargo, esto no implica que el Paebes haya alcanzado su formato definitivo debido a iniciativas recientes.

El Exeims-BC tiene un propósito claro al influir directamente en la vida estudiantil al emitir diagnósticos y determinar selección e ingreso a la educación media superior. En contraste, el Paebes y Paebes-Alfa parecen impactar más en directivos y personal docente al vincular el desempeño estudiantil con recursos financieros adicionales (Afonso, 2009; Orfield y Wald, 2000; Santos et al., 2021). Aunque hay poca evidencia de que los resultados del Paebes hayan orientado las políticas educativas hacia la mejora del aprendizaje estudiantil.

En cuanto a la calidad de la educación basada en los resultados de evaluaciones estandarizadas, esto constituye un foco de crítica entre los investigadores ya que no parece haber una comprensión efectiva de que los datos generados por los exámenes correspondan a una educación de calidad (Afonso, 2009; Klein, 2019; Pino et al., 2016). Tampoco parece haber oportunidades de participación del personal docente de diferentes materias en la elaboración de las pruebas y en la política de evaluación en su conjunto (Santos et al., 2021).

Se destaca que los resultados del Paebes, gestionada por el CAEd, se ponen a disposición del personal docente de forma restringida a través de la plataforma del Sistema de Gestión Escolar del Estado, lo que dificulta el seguimiento y acceso a los datos (Brasil, 2011) de estas evaluaciones por parte de los padres, estudiantes e investigadores. Esto se replica en Exeims-BC, donde solo un público específico dentro de las unidades escolares puede acceder a los resultados. La falta de conocimiento sobre estos resultados y su propósito, junto con la ausencia de iniciativas para mejorar el rendimiento educativo, obstaculiza la comprensión de estos exámenes por parte de los principales interesados, es decir, estudiantes y profesores.

Las evaluaciones estandarizadas, consideradas fundamentales para evaluar el desempeño estudiantil y escolar, se centran a menudo en matemáticas y lengua materna, privilegiando ciertos conocimientos y competencias curriculares (Vanlommel y Schildkamp, 2019).

Este énfasis puede tener varias implicaciones, tales como:

a) Reforzar la jerarquía del conocimiento: al enfatizar solo algunas materias, como las matemáticas y la lengua materna, las evaluaciones estandarizadas pueden crear una jerarquía del conocimiento, sugiriendo que estas materias son más importantes que otras. Esto puede llevar a una devaluación de materias como artes, ciencias sociales y educación física.

b) Estandarización del currículo: el enfoque en las matemáticas y la lengua materna puede llevar a las escuelas a centrarse excesivamente en estas materias para garantizar que el alumnado obtenga buenos puntajes en las pruebas. Esto puede dar lugar a la estandarización del currículo, en el que la enseñanza se orienta principalmente a cumplir criterios de evaluación en detrimento de la diversificación de los contenidos educativos.

c) Presión sobre los docentes y las escuelas: la importancia otorgada a algunas materias en las evaluaciones estandarizadas puede generar una presión adicional sobre los docentes y las escuelas, llevándolos a enfocarse intensamente en estos temas, a menudo descuidando otras materias igualmente valiosas para el aprendizaje del alumnado.

d) Desprecio por la diversidad de habilidades: no todo el alumnado tiene las mismas habilidades e intereses. Al priorizar solo las matemáticas y la lengua materna, las evaluaciones estandarizadas pueden no reflejar adecuadamente la diversidad de talentos y competencias del alumnado, sin tener en cuenta materias en las que algunos estudiantes puedan sobresalir.

En el caso de Paebs (Brasil) y Exeims-BC (México), estos sistemas de evaluación enfrentan presiones y desafíos similares. Las políticas basadas en evaluaciones educativas en gran escala pueden impactar ajustes y reducción curricular de acuerdo con las materias evaluadas, además del adiestramiento al alumnado en las pruebas estandarizadas conocido como *teaching to the test* (Falabella y De La Vega, 2016). Los formuladores de políticas y educadores deben considerar estos desafíos al desarrollar sistemas de evaluación y planes de estudio, promoviendo una educación equitativa que valore todas las materias y considere las necesidades individuales del alumnado y sus contextos.

Las limitaciones de este estudio incluyen la dificultad para acceder a información en línea, reduciendo el alcance de las fuentes. Además, al tratarse de dos países con un sistema político de estados federados, resulta particularmente difícil clasificarlos desde el punto de vista de los regímenes administrativos, ya que sus propios estados pueden dar lugar a reformas divergentes (Vergier et al., 2019). Sin embargo, esto impulsa la necesidad de diálogos directos con las Secretarías de Educación de Espírito Santo y Baja California para comprender las decisiones basadas en los resultados de Paebs y Exeims-BC.

Agradecimientos

Agradecemos a la *Fundação de Amparo à Pesquisa e Inovação do Espírito Santo* (FAPES) por financiar el proyecto *Avaliação estandarizada nos estados do Espírito Santo/Brasil e Baja California/México: dilemas e tensões do Paebs e do Exeims-BC*, proceso número 11367/2021, y para la beca Postdoctoral de PROFIX Circular nº 15/2022, Protocolo: 51005.788.17880.

Referencias

- Afonso, A. J. (2009). Nem tudo o que conta em educação é mensurável ou comparável: crítica à accountability baseada em testes estandarizados e rankings escolares. *Revista Lusófona de Educação*, 13(13), 13-29. <https://revistas.ucm.es/index.php/RCED/article/view/61865/45644565528415>.
- Álvarez-López, G., y Matarranz, M. (2020). Calidad y evaluación como tendencias globales en política educativa: estudio comparado de agencias nacionales de evaluación en educación obligatoria en Europa. *Revista Complutense de Educación*, 31(1), 83-93. <https://doi.org/10.5209/rced.61865>.
- Baja California. Sistema Educativo Estatal. (2012). *La evaluación educativa en Baja California: construcción de una cultura para mejorar*. Baja California: Dirección de Evaluación Educativa. https://www.educacionbc.edu.mx/departamentos/evaluacion/LibroEvaluaciones/ESTADO_DEL_ARTE_Evaluacion_TODO.pdf.
- Ball, S. J. (2016). Subjectivity as a site of struggle: refusing neoliberalism? *British Journal of Sociology of Education*, 37(8), 1129-1146. <https://doi.org/10.1080/01425692.2015.1044072>.
- Bauer, A., Alavarse, O. M., y Oliveira, R. P. (2015). Avaliações em larga escala: uma sistematização do debate. *Educação & Pesquisa*, 41, 1367-1382. <https://doi.org/10.1590/S1517-9702201508144607>.
- Beaver, J., y Weinbaum, E. H. (2015). State test data and school improvement efforts. *Educational Policy*, 29(3), 478-503. <https://doi.org/10.1177/0895904813510774>.
- Brasil. (2011). *Lei no 12.527/2011 - Lei de Acesso à Informação (LAI)*. http://www.planalto.gov.br/ccivil_03/_ato2011-2014/2011/lei/l12527.htm.
- Brown, C., y Zhang, D. (2017). How can school leaders establish evidence-informed schools: an analysis of the effectiveness of potential school policy levers. *Educational Management Administration & Leadership*, 45(3), 382-401. <https://doi.org/10.1177/1741143215617946>.

- Caballero, A., Manso, J., Matarranz, M., y Valle, J. M. (2016). Investigación en Educación Comparada: pistas para investigadores noveles. *Revista Latinoamericana de Educación Comparada: RELEC*, 7(9), 39-56.
- Carrasco, A., y Gunter, H. M. (2019). The “private” in the privatisation of schools: the case of Chile. *Educational Review*, 71(1), 67-80. <https://doi.org/10.1080/00131911.2019.1522035>.
- Caruso, M. (2011). Teoría y metodología en estudios comparados: la justificación de un plus valor y el abordaje de la globosfera. *Revista Latinoamericana de Educación Comparada: RELEC*, 2(2), 8-9.
- COBACHBC. Colegio de Bachilleres del Estado de Baja California. (2020). *Tabla de Especificaciones para EXEIMS*. Baja California, México.
- Côco, D., y Gontijo, C. M. M. (2017). Avaliação externa nas classes de alfabetização no Espírito Santo. *Pro-Posições*, 28(1), 63-87. <https://doi.org/10.1590/1980-6248-2016-0011>.
- Contreras, J., y Torres, Á. (2023). Las evaluaciones educacionales estandarizadas desde la experiencia de los actores. *Educação & Pesquisa*, 49, 1-19. <https://doi.org/10.1590/S1678-4634202349248451es>.
- Creswell, J., y Clark, V. P. (2013). *Pesquisa de métodos mistos*. Porto Alegre: Penso.
- Cunha, E. C. S., y Müller, E. R. (2018). Avaliações em larga escala: uma tentativa de controle, regulação, captura e padronização do cotidiano escolar. *Cadernos da Fucamp*, 17(29), 143-163.
- De la Vega, L., y Picazo, M. I. (2016). La accountability de la calidad y equidad de una política educativa: el caso de la subvención escolar preferencial para la infancia vulnerable en Chile. *Revista del CLAD Reforma y Democracia*, (65), 193-224, 2016.
- Ebbeler, J., Poortman, C., Schildkamp, K., y Pieters, J. (2017). The effects of a data use intervention on educators’ satisfaction and data literacy. *Educational Assessment, Evaluation and Accountability*, 29(1), 83-105. <http://dx.doi.org/10.1007/s11092-016-9251-z>.
- Elacqua, G., Martínez, M., Santos, H., y Urbina, D. (2016). Short-run effects of accountability pressures on teacher policies and practices in the voucher system in Santiago, Chile. *School Effectiveness and School Improvement*, 27(3), 385-405. <https://doi.org/10.1080/09243453.2015.1086383>.
- Espírito Santo. Secretaria de Estado da Educação. (2008). *Plano Estratégico Nova Escola*. Vitória: Sedu.
- Espírito Santo. Secretaria de Estado da Educação. (2022). *Portaria N° 138-R, de 23 de junho de 2022*. *Diário Oficial do Estado*, Vitória.
- Espírito Santo. Secretaria de Estado da Educação. (2023). *Guia informativo avaliações externas no ES*. Vitória: Sedu. <https://sedu.es.gov.br/Media/sedu/pdf%20e%20Arquivos/GUIA%20INFORMATIVO%20-%20AVALIA%C3%87%C3%95ES%20EXTERNAS%20ES%202023.pdf>.
- Falabella, A., y De La Vega, L. (2016). Políticas de responsabilización por desempeño escolar: un debate a partir de la literatura internacional y el caso chileno. *Estudios Pedagógicos*, 42(2), 395-413. <http://dx.doi.org/10.4067/S0718-07052016000200023>.
- Ferrer, F. J. (2002). *La educación comparada actual*. Barcelona: Ariel Educación.
- Bauer, A., Gatti, B. A., y Tavares, M. R. (2013). *Vinte e cinco anos de avaliação de sistemas educacionais no Brasil: origem e pressupostos*. Florianópolis: Insular.
- Gajardo, M. (1999). *Reformas Educativas en América Latina. Balance de una década*. Santiago: Chile.

- García Garrido, J. (1996). *Fundamentos de educación comparada*. Madrid: Dykinson.
- García Garrido, J., García Ruíz, M. J. y Gavari Starky, E. (2005). *La Educación Comparada en tiempos de globalización*. Madrid: UNED.
- Gentili, P., y Silva, T. T. (1994). *Neoliberalismo, qualidade total e educação*. Petrópolis: Vozes.
- Ginzburg, C. (1989). *Mitos, emblemas, sinais: morfologia e história*. São Paulo: Companhia das Letras.
- Klein, R. (2019). Uma solução para a divergência de diferentes padrões no SAEB. *Ensaio: Avaliação e Políticas Públicas em Educação*, 27(103), 229-249. <https://doi.org/10.1590/S0104-40362019002702060>.
- Libâneo, J. C., y Freitas, R. M. da M. (2018). *Políticas educacionais neoliberais e escola pública: uma qualidade restrita de educação escolar*. Goiânia: Editora Espaço Acadêmico.
- Loeb, S., y Byun, E. (2019). Testing, accountability, and school improvement. *Annals of the American Academy of Political and Social Science*, 683(1), 94-109. <https://doi.org/10.1177/0002716219839929>.
- Machado, C., Alavarse, O. M., y Arcas, P. H. (2015). Sistemas estaduais de avaliação: interfaces com qualidade e gestão da educação. *RBPAAE*, 31(3), 667-680. <https://doi.org/10.21573/vol31n32015.63800>.
- Marques, R., Stieg, R., y Santos, W. (2020). Exames estandarizados: análise dos modelos e das teorias na produção acadêmica. *Revista Meta: Avaliação*, 12(34), 1-27. <http://dx.doi.org/10.22347/2175-2753v12i34.2342>.
- Menegão, R. de C. S. G. (2016). Os impactos da avaliação em larga escala nos currículos escolares. *Práxis Educativa*, 11(3), 641-656. <https://doi.org/10.5212/PraxEduc.v.11i3.0007>.
- México. Baja California. (2022). *Objetivo general*. Secretaría de Educación del Estado de Baja California. <https://www.educacionbc.edu.mx/index.php/vision-y-objetivo-general/>.
- Míguez, D. P. (2023). Factores asociados al desempeño entre estudiantes de bajo estatus sociocultural en Brasil, Chile y Argentina. *Revista Brasileira de Educação*, 28, 1-24.
- Monarca, H. (2020). *Evaluaciones externas: mecanismos para la configuración de representaciones y prácticas en educación*. Madrid: UAM.
- Mons, N. (2009). Effets théoriques et réels des politiques d'évaluation standardisée. *Revue Française de Pédagogie*, (169), 99-140. <https://doi.org/10.4000/rfp.1531>.
- Jiménez-Moreno, J. A. (2016). El papel de la evaluación a gran escala como política de rendición de cuentas en el sistema educativo mexicano. *Revista Iberoamericana de Evaluación Educativa*, 9(1), 109-126. <https://doi.org/10.15366/riee2016.9.1.007>.
- Navarrete Cazales, Z., y Rojas Moreno, I. (2022). Estudio comparado sobre políticas para la educación secundaria en América Latina y el Caribe. *RBEC: Revista Brasileira de Educação Comparada*, 4, 1-20.
- Orfield, G., y Wald, J. (2000). Testing, testing: the high-stakes testing mania hurts poor and minority students the most. *The Nation*, 270(22), 38-40.
- Ozga, J. (2020). The politics of accountability. *Journal of Educational Change*, 21, 19-35. <https://doi.org/10.1007/s10833-019-09354-2>.
- Pino, M., Oyarzún, G., y Salinas, I. (2016). Crítica a la rendición de cuentas: narrativa de resistencia al sistema de evaluación en Chile. *Cadernos Cedes*, 36(100), 337-354. <https://doi.org/10.1590/CC0101-32622016171362>.
- Pontes Junior, J. A. de F., Osti, A., y Vidal, E. M. (Orgs.), (2021). *Avaliações estaduais: nascimento, vida e protagonismo*. Brasília, ANPAE.

- Ramos-Zincke, C. (2018). Dispositivo de evaluación y gubernamentalidad del sistema educacional: entretejimiento de ciencia social y poder. *Cinta de Moebio*, 61, 41-55. <https://doi.org/10.4067/S0717-554X2018000100041>.
- Santos, W., Soares, D. J. M., Soares, T. E. A., y Stieg, R. (2021). Paebes TRI e seus usos em escolas de ensino médio da Secretaria de Estado da Educação do Espírito Santo: qual o lugar do trabalho dos pedagogos?. En J. A. F. Pontes Junior, A. Osti, y E. M. Vidal (Eds.), *Avaliações estaduais: nascimento, vida e protagonismo* (pp. 156-181). Brasília: ANPAE.
- Soares, D. J. M., Soares, T. E. A., y Santos, W. (2022). Análise do indicador de desenvolvimento das escolas estaduais do Espírito Santo. *Educação & Sociedade*, 43, 1-20. <https://doi.org/10.1590/ES.256722>.
- Smith, W. (2017). National testing policies and educator-based testing for accountability. *OECD Journal: Economic Studies*, 2016(1), 131-149. https://doi.org/10.1787/eco_studies-2016-5jgljxftj4r3.
- Suspitsyna, T. (2010). Accountability in American education as a rhetoric and a technology of governmentality. *Journal of Education Policy*, 25(5), 567-586. <https://doi.org/10.1080/02680930903548411>.
- Sussman, J., y Wilson, M. (2019). The use and validity of standardized achievement tests for evaluating new curricular interventions in mathematics and science. *American Journal of Evaluation*, 40(2), 190-213. <https://doi.org/10.1177/1098214018767313>.
- Tolo, A., Lillejord, S., Petour, M. T. F., y Hopfenbeck, T. N. (2020). Intelligent accountability in schools: a study of how school leaders work with the implementation of assessment for learning. *Journal of Educational Change*, 21(1), 59-82. <https://doi.org/10.1007/s10833-019-09359-x>.
- Vanlommel, K., y Schildkamp, K. (2019). How do teachers make sense of data in the context of high stakes decision making? *American Educational Research Journal*, 56(3), 792-821. <https://doi.org/10.3102/0002831218803891>.
- Verger, A., Fontdevila, C., y Zancajo, A. (2017). Multiple paths towards education privatization in a globalizing world: a cultural political economy review. *Journal of Education Policy*, 32(6), 757-787. <https://doi.org/10.1080/02680939.2017.1318453>.
- Verger, A., Fontdevila, C., y Parcerisa, L. (2019). Reforming governance through policy instruments: how and to what extent standards, tests and accountability in education spread worldwide. *Discourse: Studies in the Cultural Politics of Education*, 40(2), 248-270, <https://doi.org/10.1080/01596306.2019.1569882>.



Este trabajo está sujeto a una licencia de Reconocimiento 4.0 Internacional Creative Commons (CC BY 4.0).