



Procesos Evaluativos y Pruebas Estandarizadas. ¿Son compatibles si buscamos la calidad en la educación?

Assessment Processes and Standardized Tests. Are they compatible if we look for quality in education?

Marcela Silvana Muñoz Lira¹  , Julio Andrés Bruna Gaete²  

RESUMEN

¿Existe disyuntiva entre los procesos evaluativos y las pruebas estandarizadas?, ¿son compatibles si buscamos, en estos términos, la calidad en la educación? En la literatura existen términos sobre calidad vinculados a la educación, no obstante, calidad como puntaje en una prueba estandarizada. Por otro lado, tratan la escuela desde un punto de vista del mundo empresarial, donde el éxito lo atribuyen al rendimiento académico, confundiendo calidad en la educación con condiciones estructurales y materiales necesarias. La importancia del presente artículo radica en exponer esta disyuntiva, que se abordó teóricamente, presentando una investigación de tipo indagatoria que permitió reflexionar sobre esta correlación. Los resultados mostraron una distancia latente entre los procesos evaluativos y las pruebas estandarizadas, las cuales se orientan hacia evaluaciones del tipo sumativas. Se concluye que la educación y, en particular, la evaluación, se ha transformado en un negocio lucrativo cuyo propósito es efectuar comparaciones entre individuos con determinadas y, en variadas ocasiones, características únicas, midiendo al estudiantado sin que haya un valor en el objeto a medir.

Palabras clave: calidad, docente, educación, estudiante, evaluación de la educación.

Clasificación JEL: I2, I20, I21, I22

Recibido: 10-02-2025

Revisado: 25-04-2025

Aceptado: 15-06-2025

Publicado: 04-07-2025

Editor: Carlos Alberto Gómez Cano 

¹Pontificia Universidad Católica de Valparaíso. Valparaíso, Chile.

²Universidad Andrés Bello. Viña del Mar, Chile.

Citar como: Muñoz, M. y Bruna, J. (2024). Procesos Evaluativos y Pruebas Estandarizadas. ¿Son compatibles si buscamos la calidad en la educación? Región Científica, 3(1), 2024204. <https://doi.org/10.58763/rc2024204>

INTRODUCCIÓN

Ser docente en Chile, y en Latinoamérica, requiere más amor que vocación; términos que no son cuantificables. No obstante, vocación es un término que puede causar confusión. En ocasiones, es atribuible a mártir por ser un término que se ocupa en el sacerdocio, ergo es un juicio de valor sobre el servicio que ejerce un docente. Sin embargo, ¿un docente puede otorgar distintos grados de amor según el nivel donde se desempeñe? Por ejemplo, López Arrillaga (2019) puntualiza que el desempeño del profesor primario debe estar guiado por una pedagogía sustentada en sentimientos y



atención a los procesos afectivos y el desarrollo propio de la personalidad de los estudiantes. Sin embargo, su tesis está enfocada en el docente de educación primaria, no considerando los niveles que lo suceden. Sobre el término “mártir”, la Real Academia Española lo define como quien soporta sufrimientos, injusticias o privaciones, como causa de la acción de algo o de terceros. Además, señala que esta condición se realiza si sobre las consecuencias actúa la vocación u otros procesos que valoriza la actuación de la persona, con énfasis en los aspectos religiosos.

A raíz de los textos, artículos y publicaciones que existen en la red, sobre el cuestionamiento recién planteado, surge una serie de respuestas a muchas preguntas que han nacido a lo largo de la trayectoria profesional de muchos y muchas docentes (Angulo, 2020; Angulo, 2019; Casassus, 2007; Espinosa, 2020; Espinoza, 2017; Nichols y Berliner, 2007). Sobre lo anterior, basta ingresar a un grupo de profesores en una red social como Facebook para leer la cantidad de interrogantes que los aquejan. Sin embargo, ahora se convoca a situarse en la siguiente pregunta y cuestionamiento: Procesos evaluativos y pruebas estandarizadas, ¿son compatibles si buscamos la calidad en la educación?

Concretamente en Chile, desde hace un tiempo considerable (creación del SIMCE año 2012) ha nacido y permanecido una insistente inclinación, de parte de directivos de colegios y políticos en el Ministerio de Educación, sobre elaborar y aplicar instrumentos de evaluación que no cumplen con la base de la educación, ni representa el amor docente o el interés interior por dedicarse a esta forma de vida (la docencia). La educación debe estar destinada a la formación, a promover el desarrollo de las capacidades intelectuales y afectivas, a preparar a las futuras generaciones para hacer el mejor uso posible del acervo de conocimiento y experiencias compartidas, así como aportar al mejoramiento social. En este sentido, Casassus (2007) escribió un artículo inspirador y revelador que muestra cómo las pruebas estandarizadas referidas a normas merman la finalidad de la educación, fomentado la desigualdad y disminuyendo la calidad de la educación, cuestionándose, además, lo siguiente:

¿Qué es lo que ha ocurrido que se abandonó la tradición de educar para la formación humana, de preocuparse por los alumnos, de educar para mejorar la sociedad? ¿Por qué se ha obligado a los docentes a abandonar el entusiasmo por su trabajo, haciéndolos entrar en el juego de las presiones y amenazas? (p. 78). La segunda pregunta cala fuertemente, pues expresa aquellos sentimientos reprimidos que tal vez el profesorado ha guardado por años; docentes que ingresaron al sistema educativo con un afán loable para mejorar el mundo, para contribuir en la formación de las futuras generaciones. Sin embargo, estos se encuentran con presiones y amenazas dentro del mismo sistema escolar al momento de cambiar las reglas tradicionalistas de las clases o simplemente cuando cuestionó por qué se ha abandonado la tradición humana de educar para mejorar la sociedad; esto a la luz de la primera pregunta que plantea Casassus (2007).

En Chile, desde 1968, existe un sistema educativo basado en la evaluación externa y la doble funcionalidad de la generación de datos valiosos para la toma de decisiones de los actores educativos. Su objetivo general declarado fue sustentado en dos baluartes de la educación; en los discursos del mainstream académico y político, la calidad y la equidad. O sea, a partir de una metodología cuantitativa y una epistemología positivista se generarían los datos necesarios para evaluar los aprendizajes.

El Sistema de Evaluación de la Agencia de Calidad de la Educación (SIMCE), existe desde el año 2012; su propósito es evaluar los resultados de aprendizajes de las instituciones educativas a través de un proceso de medición y comparación que se aplica al estudiantado de todo el país en varios cursos del sistema escolar, en las áreas de lenguaje y matemática (2do básico, 4to básico, 6to básico, 8vo básico, II año medio, y recientemente se ha agregado III año medio en el área de inglés). En su página web señalan que recogen información sobre docentes, estudiantes, padres y apoderados a través de un cuestionario (una medida cuantificable). Esta información es utilizada para analizar contextualmente los resultados del estudiantado de Chile, pero ¿cómo conciliamos los procesos educativos con las pruebas estandarizadas si el proceso educativo no es una medida, no es cuantificable y los resultados del aprendizaje no son un número?

Como se ha mencionado, existen términos ambiguos que han sido empleados en educación y que, por ignorancia o prudencia, no se cuestionan (tal vez) ni se profundiza en la praxis cotidiana sobre origen o naturaleza de su uso. A continuación, nos preguntamos ¿qué entendemos por calidad de la educación?

Diferentes autores han hecho un análisis de esta frase, pero se intenta en la reflexión no separar sus términos: calidad y educación. Por ejemplo, a principios del siglo XX se debatió acerca de cómo se debía definir qué es la inteligencia. Concluyeron que la inteligencia sería lo que miden los test de inteligencia, símil de lo que ocurre con calidad de la educación, pues el aprendizaje se consideraría equivalente a un puntaje en una prueba estandarizada (Angulo, 2020; Angulo, 2019; Casassus, 2007; Espinoza, 2017; Nichols y Berliner, 2007).

Pero inicios de la década de los 80, como menciona Casassus (2007), ocurrió un giro en la política de educación a nivel mundial que no tardó en llegar a América Latina. Este giro se concentró adentro del sistema, pasando de un enfoque centrado en la cantidad hacia uno centrado en la calidad, pero también con la intervención de agencias estatales en el diseño de indicadores de medición (Burstein, 1988). Sin embargo, no explica que se entienda por calidad de la educación. Por otro lado, Espinosa (2020) menciona que desde los contextos de empresa procede el término “calidad de la educación”, que actúa como la pócima curativa a la crisis actual en el mundo educativo. Sin embargo, la solución radica en relegar la educación al docente transformando su pedagogía en una función administrativa. Por ende, Mejía (2006) indica que los resultados de los establecimientos educacionales se comportan como el mundo de la fábrica donde deben mostrarlos como productos.

Por su parte, Lovaglio señala que en los países de América Latina y el Caribe existen dos interpretaciones concretas sobre “calidad de la educación” (2016, p. 6). Por un lado, se entiende como la base de la convivencia y la democracia, relevando la importancia de las dimensiones ciudadanas, cívicas y valóricas. Y, por otro lado, se relaciona con los efectos socioeconómicos de la educación, que son limitadas por factores como el aporte al crecimiento económico, el acceso al empleo y la integración social; elementos que garantizan el acceso de la población a la escuela donde todos y todas puedan acceder a las competencias, conocimientos, habilidades y valores que los modelos educativos actuales establecen como indicadores de calidad prometida.

En otro punto, Lovaglio (2016) señala que: “cuando se hablamos de “calidad”, la mayor parte de las veces adjudicado a las evaluaciones estandarizadas, resulta un tanto excesivo ya que hay más en la calidad educativa que no puede ser adjudicado por las evaluaciones estandarizadas” (p.13). De lo anterior, es posible plantear dos perspectivas. Una es la concepción “tecnocrática”, donde se vincula la baja calidad en educación con la ineficiencia que brinda el sistema educativo y que las acciones para mejorarla dependen de decisiones de carácter técnico y pedagógico (Filmus, 1997). Estas decisiones impactan en el carácter productivo de los actores y de las instituciones, pero también su posición secundaria con respecto a actores externos (Aydarova, 2021; Cruickshank, 2019; Mehta, 2013).

La segunda concepción se orienta desde la perspectiva de mercado, donde la calidad de la educación se somete a las leyes de oferta y demanda, esto crea situaciones de competencia entre los establecimientos escolares. Además, supone la supeditación a los movimientos del mercado, la penetración de factores económicos y la sustitución de los aspectos propiamente educativos en la toma de decisiones (Khoshtaria *et al.*, 2020; Williamson, 2021; Williamson *et al.*, 2020).

En ambas concepciones, la preocupación por la educación se parcializa y se tiende a elitizarla porque la calidad en la educación se transforma en una calidad privativa ya que es para pocos; hecho amparado por medidas de selectividad y una cultura instaurada de la deseabilidad del estudiante de élite (Lorbeer, 2020; Telling, 2020). Estas perspectivas absolutizan la relevancia de los factores endógenos que se producen en el entorno y aquellos elementos controlados de manera directa por cada una de las instituciones educativas. El riesgo en ambas perspectivas es no considerar que para que exista una calidad en la educación, se debe atender a aquellos factores que están fuertemente determinados por los contextos socioeconómicos y por las políticas educativas. En consecuencia, se debe poner atención a variadas dimensiones o enfoques que son complementarios entre sí.

Casassus (2007) menciona varias de las iniciativas diseñadas para el mejoramiento de la calidad del proceso educativo, como el aumento del fondo de horas-clase; extensión en el calendario curricular; descentralizar los currículos; estandarizar y centralizar la evaluación; mayor privatización y competencia regional entre centros educativos, entre otras. De manera que el autor señala que estos serían los principales elementos a considerar por líderes escolares y administrativos, quienes también son un punto focal en las iniciativas para “elevar” la calidad educativa.

En este sentido, Lovaglio (2016) indica que la calidad es un concepto político que está relacionado a la idea de educación como un derecho y proceso que está al servicio de construir una sociedad más justa y democrática donde cada niño y niña asista a la escuela, al ser el contexto por excelencia de las vivencias significativas en las primeras etapas. Sin embargo, la historia transcurrida hasta la fecha muestra que varias de estas políticas favorecieron la gestión del sistema, no de la educación (Salazar y Rifo, 2020). Estas corrientes no se generaron desde un examen profundo de la filosofía, la psicología, la pedagogía o la sociología, sino en la economía y los mecanismos políticos de control (San Martín *et al.*, 2015; Vázquez, 2015). Lovaglio señala, además, que la Unesco plantea el concepto de la calidad en la educación como un abanico de diversos significados que, con frecuencia, no coinciden entre los distintos actores porque este concepto “implica un juicio de valor respecto del tipo de educación que se quiere para formar un ideal de persona y de sociedad” (p. 5). Esto se debe porque la definición del término “calidad educativa”

se transforma y se adecúa según el periodo, sociedad o grupo de individuos que en determinado momento estén desarrollándola.

Adicionalmente, Blanco (2007) menciona que con frecuencia el término “calidad educativa” se asimila con eficiencia y eficacia, donde se considera la educación como un producto y como un servicio propio de lo que se consideraría un mercado educativo, que debe satisfacer las necesidades de los usuarios. Sin embargo, cuando se habla de educación, esta se debe fundamentar en un conjunto de valores y concepciones que se contraponen con el juicio de valor que se tiene acerca de la calidad de la misma (Lovaglio, 2016).

Al analizar el fenómeno de la calidad en la educación, primero se debe aludir a la eficacia, pues existen elementos asociados al resultado del proceso, como objetivos de aprendizaje o desempeños, que están contenidos en los programas y deben ser logrados. Además, se entiende que la calidad se basa en una contraparte física, que representa el entorno adecuado para la conducción del proceso de enseñanza aprendizaje, en el cual interactúan estudiantes y un claustro preparado para enseñar y que sustenta su actuación en estrategias didácticas y pedagógicas adecuadas. En términos simples, la calidad de la educación debe considerar aspectos relacionados con la eficacia y la eficiencia, pero también debe respetar los derechos de toda la comunidad educativa para que sea un proceso relevante, pertinente y equitativo.

METODOLOGÍA

Enfoque y diseño

Esta investigación fue de tipo reflexiva, ya que se exponen resultados de un barrido investigativo terminado, siendo un artículo creado desde el punto de vista analítico y crítico de los investigadores. El mismo se sustenta en un paradigma crítico reflexivo, pues su foco fue cualitativo y se consideraron las connotaciones del tipo interpretativas necesarias para profundizar en la presente investigación. Esta investigación surgió como resultado de un curso del magíster en Educación, con mención en Evaluación Educativa, dictada por la Pontificia Universidad Católica de Valparaíso, y ejecutada por el Dr. Félix Ángulo Rasco. El curso, de tipo cualitativo, buscaba potenciar la reflexión crítica sobre la temática de la Evaluación y Estándares de la Educación.

Estrategia de búsqueda

Los materiales usados fueron artículos indexados en Scopus, Springer Link, latindex, Dialnet, *Scielo*, etc. Primero se realizó una búsqueda en español, por la lengua nativa de los autores, y posteriormente en inglés, debido a su alto impacto de las publicaciones a nivel internacional.

Se empleó el uso palabras claves en buscadores como Google Académico, con filtros de tiempo, relevancia, idioma y tipos de artículos, asimismo en bases de datos, como: *A+Education Database*; *Education Full Text*; ERIC; *Education Research Complete*; Tesauro de la Unesco, etc. En la búsqueda se utilizaron palabras claves en español y en inglés, entre ellas: “calidad en educación”, “quality in education”; “educational quality” “evaluación de la educación”; “educational assessment”; “estudiante”; “student”; “docente”; “teacher”; “calidad de la educación”; “estudiante de primaria”; “estudiante de secundaria”; “docente de primaria”; “docente de secundaria”; “docente especializado”; “school teacher”.

Los criterios de selección se desglosaron en una serie de pasos que permitieron eliminar artículos con una antigüedad mayor a 10 años y almacenar aquellos relevantes en una base de datos complementaria. Se seleccionaron investigaciones con ambos enfoques puros (cual-cuan) y en mixtura. Los textos seleccionados fueron insertados con sus metadatos en una Hoja de Cálculo para proceder a ordenarlos, según su enfoque metodológico.

De esta manera, se persiguió un procedimiento similar a los realizados en la metodología de *scoping review* para la exploración de temáticos o primeras aproximaciones a un objeto (Ruiz y Petrova, 2019; Westphal *et al.*, 2021), pero sin llegar al grado de rigurosidad propuesto en la declaración PRISMA, pues el estudio se dirigió más hacia la comprensión y menos a la generalización. El propósito fue coordinar los análisis y lograr una síntesis de conocimientos sobre la pregunta principal de estudio: Procesos Evaluativos y Pruebas Estandarizadas, ¿son compatibles si buscamos la calidad en la educación?

RESULTADOS

En esta etapa resultó necesario presentar los hallazgos en dos partes. Primero en el punto, como Proceso de evaluación; luego, como Pruebas estandarizadas. El motivo radica en lograr realizar un contraste según la información hallada. Así, la discusión resultará de forma clara, ordenada y precisa, logrando una consonancia con el objetivo planteado.

Proceso de evaluación

Este proceso tuvo su inicio en gran parte de los países de América Latina y el Caribe a principio de los años 90 (Gysling, 2016; Lovaglio, 2016; Rodrigo, 2019). Con la aspiración de modificar los diseños previos de la organización escolar e institucional del sistema educativo, en su mayoría, los países crearon sistemas nacionales de evaluación y de medición de la calidad en la educación. Esto generó la elaboración de mecanismos regionales de medición de resultados y se sumaron a instrumentos de medición internacionales como la prueba PISA (Programa Internacional de Evaluación de los Alumnos), TIMMS (*Trends in International Mathematics and Science Study*) o PERCE, SERCE y TERCE, que son pruebas de Estudio Regional Comparativo y Explicativo que se realizan cada 6 años en quince países de América Latina y el Caribe. Estas pruebas estandarizadas ofrecen datos sobre diversos aspectos del fenómeno educativo (Delprato, 2019; Delprato y Akyeampong, 2019; Fernández *et al.*, 2019).

Sobre estas evaluaciones, a pesar de la variedad de datos sociodemográficos que aportan, Lovaglio (2016) menciona que son riesgosas, ya que cualquier sistema de evaluación de resultados debe pretender la globalidad y no reducir la tarea educativa a la mera forma instructiva. Además, estas formas de evaluar la educación realizan comparaciones y clasifican a los establecimientos de forma injusta sin considerar sus contextos ni particularidades (Díaz y Osuna, 2016; Ravela, 2001), lo cual suele entenderse como una debilidad en la validez transcultural de los hallazgos y requiere de complejos meta-análisis para poder emplear de forma justa los datos procesados (Fernández *et al.*, 2019; Strello *et al.*, 2021).

Varias de las fuentes consultadas critican las fallas en la lógica tras la estandarización, entre otras razones, porque su carácter predictivo no es infalible y podrían ser más útiles estos instrumentos con fines descriptivos, pues no evalúan realmente aprendizajes, sino cuántos estudiantes se encuentran en una norma establecida a priori (Berliner, 2020; Lovaglio, 2016). Esto último, a la luz de la importancia que cobra la retroalimentación en todo proceso de evaluación formativa, es relevante para comprender que la evaluación de tipo estandarizada no muestra interés por mostrar situaciones individuales, pues busca dar cuenta del sistema educativo, pero en conjunto, sin detenerse en particularidades, dado que presenta una visión global (Contreras y Zúñiga, 2017; Muñoz y Solís, 2021; Muñoz, 2020).

Esto significa que este tipo de modelos de evaluación no se preocupa por la diversidad socioeducativa, en tanto sus diseños no analizan la situación puntual de aprendizaje de cada estudiante, sino que diluyen las singularidades en datos crudos. Sin embargo, en estas particularidades cobra un papel preponderante el docente, pues su evaluación interna está sujeta a fenómenos subjetivos y se preocupa de diseñar un instrumento que se ajusta a las necesidades particulares de sus estudiantes.

Pruebas estandarizadas

Sobre las pruebas estandarizadas, diversos autores consideran que recalcan las debilidades del estudiante (Espinoza, 2017). En Chile se cuenta con el Sistema de Medición de la Calidad de la Educación (SIMCE), una prueba estandarizada referida a norma. Su propósito es determinar posiciones y establecer *ranking* entre los colegios que las responden. Este sistema educacional nacional se basa en un currículo controlado/reducido que mide los conocimientos básicos del currículo, que además ofrece la posibilidad al Estado de determinar y aplicar medidas como la intervención y cierre de establecimientos escolares (Gysling, 2016).

En este sentido, profesores, estudiantes, directivos, apoderados, es decir, la comunidad educativa en conjunto, gira en torno a esta evaluación sumativa. Literalmente, el colegio o el establecimiento educativo se paraliza cuando se avecina, cambiando (o desestimando) la planificación que elaboró el docente, obligándolo (aunque suene imprudente) a convertirse en un operador, un técnico. En este escenario, el profesorado pierde su estatus profesional, y se convierte en un ejecutador que sanciona cuando los resultados no son los esperados, o bien premie en el caso contrario o esperado. Despóticamente, transforma al docente en un ente punitivo.

¿Resultados? Consecuencias sociales, como lo explican Ángulo (2020), Casassus (2007) o Nichols y Berliner

(2007). Una consecuencia social es el aumento de la desigualdad donde colegios con bajos resultados presionan a sus profesores a no educar, sino a entrenar a sus estudiantes para responder pruebas de alternativas. En estas realidades, si las familias tienen un ingreso económico mayor invierten en clases particulares (fuera de la hora de clases o jornada escolar) para “entrenar” a sus hijos en la rendición de pruebas estandarizadas, como es el caso de la PSU (Prueba de Selección Universitaria) o PTU (Prueba de Transición Universitaria), que fue de transición entre la PSU y la PAES. Desde el año 2022 contamos con la PAES (Prueba de Acceso a la Educación Superior), creada por el Departamento de Evaluación, Medición y Registro Educativo (DEMRE) de la Universidad de Chile.

Con certeza, muchas familias suponen su nivel de ingreso según el resultado de sus hijas e hijos en las evaluaciones sumativas, sin considerar que estas evaluaciones son del tipo selección múltiple, que no validan el aprendizaje obtenido o adquirido. En este sentido, Espinoza (2017) menciona que, vista la escuela desde el mundo empresarial, el éxito académico se reduce al rendimiento (nota, calificación) en estas pruebas, lo cual se produce desde la confusión entre calidad y las condiciones materiales que envuelven al proceso y reproducen inequidades y otros fenómenos sociales similares (Redon *et al.*, 2021). Además, Berliner (2020) señala una falla fundamental, la insensibilidad de estas pruebas respecto a las cualidades de la escuela y el profesorado. En cambio, son extremadamente sensibles a factores socioeconómicos, como las clases sociales y los niveles de pobreza, de ahí la importancia de la idea anteriormente mencionada.

DISCUSIÓN

Ángulo (2020) explica que los test no suponen evaluación, pues la evaluación pone el énfasis en la calidad del servicio, por lo que su interés son los componentes del proceso y la actuación de las personas que en ellos participan. En cambio, *testing* es la aplicación de pruebas estandarizadas, las cuales discriminan entre aquellos individuos que han aprovechado las posibilidades que el sistema de enseñanza les brinda, discriminando y seleccionando entre aquellos que se han preparado para rendirla. Al referirse a los procesos formativos y sumativos, se habla de “assessment”. Angulo Rasco (2019) señala que “assessment” se refiere al impacto que ha tenido determinado servicio en los receptores (estudiantes). Dentro de un proceso de evaluación puede entrar un proceso de assessment, de manera que su jerarquía debe ser adecuadamente explorada, pero en ningún caso limitados a un fenómeno de equivalencia.

Es importante conocer, comprender y aplicar esta diferencia anglosajona, pues se abusa de algunos términos para explicar procesos de enseñanza sin detenerse a pensar si realmente es correcto lo que se quiere transmitir. Hablar de evaluación sin hacer un quiebre en el discurso cuando se pretende expresar a la comunidad educativa sobre los procesos de evaluación, el término *assessment* debería comenzar a utilizarse marcando una distinción o diferencia cuando se refiere a un proceso de evaluación o *testing* en el caso de ocupar un instrumento.

Por consiguiente, existen docentes que basan sus clases en la aplicación de test y esto lo atribuyen a un proceso evaluativo cuando están preparando a sus estudiantes para rendir una prueba estandarizada. Esta realidad, contrastada en la experiencia de los autores, sugieren las siguientes interrogantes:

¿Reflexionarán si el uso de este instrumento está centrado para denotar los procesos evaluativos del estudiante?

¿Se preguntarán si la prueba SIMCE es representativa del contenido curricular?

¿Cuestionarán si el contenido curricular nacional es representativo del contenido curricular disciplinar?

Específicamente, el SIMC genera una pobre representación de la cultura y desarrollo curricular chileno. Lo anterior se refiere al constructo a evaluar (medir), pues este cambia, es sensible al contexto, como se ha mencionado previamente. Tal como lo señalan Angulo (2020); Angulo y Redon (2023), el SIMCE sirve para estigmatizar, discriminar y controlar escuelas, docentes e individuos, si bien su propósito original fue otorgar información relevante para una mejor toma de decisiones parentales (Gysling, 2016). Asimismo, Nichols y Berliner (2007) señalan que estas pruebas clasifican a la gente, instalando una nueva forma de discriminación. En Chile, en el año 2010 aconteció el uso de un “mapa de semáforos SIMCE”, idea del ministro de Educación de ese periodo, quien propuso utilizar el sistema de clasificación de colores para ubicar colegios con respecto a la media.

Según la fundamentación de esta medida, ayudaba a sus padres a escoger un “colegio bueno” en vez de un “colegio malo”, motivando a los “colegios malos” a aumentar su matrícula porque estaban “perdiendo alumnos, a ponerse las pilas y empezar a mejorar”. En la práctica educativa esto generó efectos contraproducentes, sin considerar el progreso de las escuelas, comparando establecimientos con distintas realidades, donde el número de estudiantes

que rindieron el SIMCE no aseguraba la validez del resultado obtenido. Eso sin aludir a la estigmatización, ya que colegios con estudiantes con una alta vulnerabilidad social no fueron considerados en función del entorno geográfico, familiar ni emocional; factores exógenos a la calidad del colegio y que influyen en el resultado del test. Esta medida sí logra mostrar las desigualdades sociales y “desnudó” la educación existente en Chile.

Los resultados expuestos indicaron que, en apariencia, parecía una buena medida que dejaba que el mercado funcionara: tratando la educación como un fin lucrativo, un bien de consumo, y no un bien garantizado. Sin embargo, esta medida trascendió como un plan que, pese a sus defectos, evidenció el abismo existente entre colegios de dependencia privada, subvencionada y municipal. A continuación, se presenta una síntesis de los hallazgos analizados en las principales fuentes de este estudio de reflexión:

- Nichols y Berliner (2007) señalan que las pruebas estandarizadas fueron hechas para medir la productividad y, sin gastar dinero, hacer responsables a las escuelas y los educadores.
- Espinoza (2017) se refiere a cómo la cultura pedagógica se transformó en una cultura del mercado. Indica que los organismos internacionales fomentan la pérdida paulatina de la autonomía de los sistemas escolares, que terminan por someterse a políticas educativas mercantiles. Señala que estudiantes se transformaron en productos, padres en clientes, escuelas en empresas, transformando la educación en un producto en vez de un proceso.
- Sobre calidad de la educación, Espinoza (2017) indica que es un término empresarial de 1980, de cuño neoliberal, que está relacionado con la jerga organizacional e indicadores comunes en ese entorno (por ejemplo, eficiencia y eficacia, bajo costo, rentabilidad y excelencia).
- En este sentido, Sahlberg (2016) usó el término “*Global Educational Reform Movement*” (GERM) para describir el surgimiento de una nueva ortodoxia global en la política educativa e identificó sus principales características:
 - _ La competencia entre las escuelas, lo cual conduciría a mejores resultados.
 - _ La autonomía de las escuelas, necesaria para que compitan adecuadamente. _ Libertad para que los padres elijan escuelas para sus hijos.
 - _ Información para el público, basada en medidas comparables de rendimiento estudiantil y en un único plan de estudios nacional.
 - _ Mayor estandarización y estrechamientos del currículo para enfocarse en materias y conocimientos básicos.
 - _ Uso de prácticas de gestión corporativa, como las características clave de la nueva ortodoxia.
 Estos indicios políticos comenzaron hace más de tres décadas en las reformas educativas introducidas en los países de Estados Unidos, Chile y el Reino Unido.
- Sobre estandarización, Sahlberg (2016) señala que es demasiado estricta, reduciendo la libertad y la flexibilidad en las escuelas para promover cambios verdaderamente significativos para los jóvenes. Es decir, que limita la enseñanza y el aprendizaje. Entre más pruebas, menos libertad, creatividad, riesgo y aprendizaje.
- Lovegio (2016) señala que, aunque las pruebas no pueden rendir cuenta sobre la realidad en su totalidad, sí permiten conocer las líneas generales sobre las necesidades a considerar en el diseño de la política pública. Señala que uno de los temas educativos más preocupantes es conocer las causas tras el porqué de que unos centros educativos tengan mejores resultados cuando se proponen alcanzar sus objetivos, mientras que otros no logran los resultados esperados, incluso considerando contextos similares. Esto atañe a investigadores, docentes, directivos, políticos, en fin, a toda la comunidad que circunscribe el mundo educativo.
- Murphy (2010) también distingue la prueba PISA por su conexión con la OCDE, entendiéndola como una ventaja o desventaja comparativa pues puede ser determinada y utilizada política y económicamente. Indica que es un instrumento maniobrado que privilegia ciertos tipos de conocimientos.

Cuando se habla de pruebas estandarizadas se refiere a evaluaciones que responden al acto de medición de atributos y comparación entre participantes, así como con una hipótesis de funcionamiento. Estos *testing* no contribuyen directamente a mejorar la calidad de la educación, ni son necesariamente una plataforma para que el profesorado pueda desarrollar mejores clases; incluso no son útiles para la gestión (Casassus, 2007).

El propósito final de estas pruebas, enmascarado entre principios altruistas, es de carácter profundamente positivista y permite establecer escalafones y distinciones entre los sujetos del *pool* de participantes. Si esta noción basada en baremos, cuartiles y separación de miembros en función de números crudos se triangula con los factores psicosociales abordados previamente, se comprenden los impactos en términos de discriminación y de generación de ventajas estructurales que perpetúan el *status quo*.

En atención a estos argumentos, es necesario revisar los vigentes planteamientos de Glaser (1998), quien afirma que la estandarización de la evaluación no es más que instituir una medición arbitraria que, en definitiva, no aporta información relevante con respecto a qué saben “verdaderamente” los estudiantes o sobre la calidad del proceso docente. El principal aporte de este autor es señalar que para tal fin es vital diversificar la evaluación y triangular los resultados a partir de diferentes vías para explorar qué saben los estudiantes.

Entre estas vías resaltan los portafolios; las pruebas para evaluar desempeños y distintas experiencias asociadas a proyectos; la resolución de problemas prácticos e iniciativas semejantes. Sin embargo, los autores consideran válido destacar que incluso en estas alternativas se corre el “riesgo” de la medición como criterio, que se contrasta en la tendencia de los sistemas educativos a establecer criterios numéricos para definir la calidad de un aprendizaje y su proyección en la actuación de los estudiantes y como medida de evaluación de los profesores.

Por último, el profesorado debe seguir aprendiendo, debe conocer nuevas metodologías, sumergirse en el mundo de la didáctica desde su especialidad (Acaso, 2014; Mena, 2010). Esto supone no conformarse con lo aprendido en pregrado porque esto será la punta del iceberg; la inversión constante en la educación durante toda la vida y la formación, permanecen como las principales vías para promover la calidad educativa.

CONCLUSIONES

A raíz de la síntesis realizada, la triangulación de las experiencias docentes de los autores y la valoración general, se concluye que: si bien las diferencias entre evaluación y medición, así como su importancia en el sector educativo, cuentan con una amplia literatura dedicada al campo, la realidad educativa dista de ser la deseable. El análisis que se condujo apunta no solo a dificultades conceptuales arraigadas en el discurso y la práctica, sino severos problemas estructurales frecuentemente invisibilizados.

Estos problemas estructurales afectan a todos los actores educativos, tanto individualmente, como sector (estudiantes, docentes, administrativos, comunidades), o desde sus interrelaciones. Por tanto, el tema analizado va más allá de la concepción tradicional y su clasificación como educativo negaría las interdependencias entre los diferentes sectores de la sociedad.

Se espera que las reflexiones realizadas puedan contribuir a la discusión actual sobre el impacto de la estandarización de la evaluación, así como sus consecuencias sociales y, puntualmente, sobre la figura del docente. Los datos consultados en las fuentes, así como la valoración del estado del arte, sugieren la necesidad de profundizar en la temática mediante futuros estudios, principalmente de enfoque mixto y con atención especial a las vivencias relacionadas, de la perspectiva de múltiples actores.

REFERENCIAS

- Acaso, M. (2014). *rEDUvolution*. Santiago de Chile: Paidós.
- Angulo, F. (2019). Evaluación y estándares en educación: Legitimidad y equidad. Serie Por una Educación Inclusiva, 2. Valparaíso, Chile: EduInclusiva. https://eduinclusiva.cl/wp-content/uploads/2019/12/Serie02_13-12.pdf
- Angulo, F. (2020). Sobre/Contra el SIMCE. Propuestas para políticas inclusivas. Valparaíso, Chile: Centro de Investigación Educación Inclusiva. 9p. https://eduinclusiva.cl/wp-content/uploads/2020/05/Policy-Brief_Contra-Simce_Angulo_14-0520-1.pdf
- Angulo, J., y Redon, S. (2023). Un currículum constituyente para Chile. *Revista Brasileira de Educação*, 28, e280076. <https://doi.org/10.1590/S1413-24782023280076>
- Aydarova, E. (2021). Building a One-Dimensional Teacher: Technocratic Transformations in Teacher Education Policy Discourses. *Educational Studies*, 57(6), 670-689. <https://doi.org/10.1080/00131946.2021.1969934>
- Berliner, D. (2020). The Implications of Understanding That PISA Is Simply Another Standardized Achievement Test. En G. Fan, y T. Popkewitz, *Handbook of Education Policy Studies*, 239-244. Springer. https://doi.org/10.1007/978-981-13-8343-4_13

- Blanco, R. (2007). Educación de calidad para todos un asunto de Derecho Humano. Documento de discusión sobre políticas educativas. Santiago de Chile: Unesco: Prelac.
- Burstein, L. (1988). Educational quality indicators in the United States: Latest developments. *Studies in Educational Evaluation*, 14(1), 75-89. [https://doi.org/10.1016/0191-491X\(88\)90021-1](https://doi.org/10.1016/0191-491X(88)90021-1)
- Casassus, J. (2007). El precio de la evaluación estandarizada: la pérdida de calidad y la segmentación social. *Revista Brasileira de Política e Administração da Educação-Periódico científico editado pela ANPAE*, 23(1), 71-79. <https://seer.ufrgs.br/index.php/rbpaee/article/view/19014>
- Contreras, G., y Zúñiga, C. (2017). Concepciones de profesores sobre retroalimentación: Una revisión de la literatura. *Magis, Revista Internacional de Investigación en Educación*, 9(19), 69-90. <https://doi.org/10.11144/Javeriana.m9-19.cpsr>
- Cruikshank, J. (2019). Economic Freedom and the Harm of Adaptation: On Gadamer, Authoritarian Technocracy and the Re-Engineering of English Higher Education. *Social Epistemology*, 33(4), 337-354. <https://doi.org/10.1080/02691728.2019.1638988>
- Delprato, M. (2019). Parental education expectations and achievement for Indigenous students in Latin America: Evidence from TERCE learning survey. *International Journal of Educational Development*, 65, 10-25. <https://doi.org/10.1016/j.ijedudev.2018.12.004>
- Delprato, M., y Akyeampong, K. (2019). The effect of working on students' learning in Latin America: Evidence from the learning survey TERCE. *International Journal of Educational Development*, 70, 102086. <https://doi.org/10.1016/j.ijedudev.2019.102086>
- Díaz, K., y Osuna, C. (2016). Las evaluaciones estandarizadas del aprendizaje y la mejora de la calidad educativa. *Temas de Educación*, 22(1), 131-146. <https://revistas.userena.cl/index.php/teeducacion/article/view/741>
- Espinosa, D. (2020). *Práctica pedagógica sentipensante como resistencia ante los discursos de la calidad de la educación: la experiencia biográfica narrativa de un maestro de una institución educativa privada*. [Tesis de grado, Universidad de Antioquía]. https://bibliotecadigital.udea.edu.co/bitstream/10495/16443/5/EspinosaDeyby_2020_Pr%C3%A1cticasPedagogicasSentipensantes.pdf
- Espinoza, D. (2017). La escuela no es una empresa, ni la educación un negocio. *Praxis Pedagógica*, 21, 65-78. <https://doi.org/10.26620/uniminuto.praxis.17.21.2017.65-78>
- Fernández, R., Woitschach, P., Álvarez, M., González, A., Cuesta, M., y Muñoz, J. (2019). Homework and Academic Achievement in Latin America: A Multilevel Approach. *Frontiers in Psychology*, 10. <https://doi.org/10.3389/fpsyg.2019.00095>
- Filmus, D. (1997). La descentralización educativa en Argentina: elementos para el análisis de un proceso abierto. Coloquio Regional sobre Descentralización de la Educación en América Central, Cuba y República Dominicana, 1-14. San José, Costa Rica: Ministerio de Educación Pública.
- Glaser, R. (1998). Education for all: Access to learning and achieving usable knowledge. *Prospects*, 28(1), 5-20. <https://link.springer.com/article/10.1007/BF02737776>
- Gysling, J. (2016). The historical development of educational assessment in Chile: 1810–2014. *Assessment in Education: Principles, Policy & Practice*, 23(1), 8-25. <https://doi.org/10.1080/0969594X.2015.1046812>
- Khoshtaria, T., Datuashvili, D., y Matin, A. (2020). The impact of brand equity dimensions on university reputation: an empirical study of Georgian higher education. *Journal of Marketing for Higher Education*, 30(2), 239-255. <https://doi.org/10.1080/08841241.2020.1725955>
- López, C. (2019). La Pedagogía del Amor y la Ternura: una práctica humana del docente de educación primaria. *Revista Científica*, 4(13), 261-277. <https://orcid.org/0000-0002-2926-8508>

- Lorbeer, E. (2020). Contemporary Elitism in Higher Education in the United States: An Issue Scan. SSRN. <https://doi.org/10.2139/ssrn.3512115>
- Lovaglio, R. (2016). El papel de la Dirección de Evaluación Calidad Educativa de la provincia de Mendoza en el sistema educativo desde 1992. Perspectivas actuales en la investigación en ciencias sociales: problemáticas, enfoques epistemológicos y abordajes teórico-metodológicos, 1-27. Mendoza: Universidad de UNCuyo. <https://bdigital.uncu.edu.ar/9940>
- Mehta, J. (2013). The Penetration of Technocratic Logic into the Educational Field: Rationalizing Schooling from the Progressives to the Present. *Teachers College Record: The Voice of Scholarship in Education*, 115(5). <https://doi.org/10.1177/016146811311500507>
- Mejía, M. (2006). Educación(es) en la(s) globalización(es) I: Entre el pensamiento único y la nueva crítica. Bogotá: Ediciones desde abajo.
- Mena, A. (2010). Acerca de la importancia de la Didáctica de la Matemática para nuestro país. Valparaíso: Facultad de Ciencias Instituto de Matemáticas.
- Muñoz, E. y Solís, B. (2021). Enfoque cualitativo y cuantitativo de la evaluación formativa. *ReHuSo*, 6(3), 1-13. <https://revistas.utm.edu.ec/index.php/Rehuso/article/view/1684>
- Muñoz, M. (2020). Análisis de las prácticas declaradas de retroalimentación en matemáticas, en el contexto de la evaluación, por docentes chilenos. *Perspectiva Educacional. Formación de Profesores*, 59(2), 111-135. <http://dx.doi.org/10.4151/07189729-vol.59-iss.2-art.1062>
- Murphy, S. (2010). El arraste de PISA: intertidumbre, influencia e ignorancia. *Revista Interamericana de Educación para la Democracia*, 3(1), 30-48. <https://scholarworks.iu.edu/journals/index.php/ried/article/viewFile/616/1070>
- Nichols, S., y Berliner, D. (2007). *Collateral Damage. How high-stakes testing corrupts america's schools*. Massachusetts: Harvard Education Press Cambridge.
- Popkewitz, T. (2013). PISA: números, estandarización de la conducta y la alquimia de las materias escolares. *Revista de Curriculum y formación del profesorado*, 17(2), 47-64. <http://www.ugr.es/local/recfpro/rev172ART3.pdf>
- Ravela, P. (2001). ¿Cómo presentan sus resultados los sistemas nacionales de evaluación educativa en América Latina? Lima: Preal.
- Redon, S., Vallejos, N., y Angulo, J.. (2021). Education for Citizenship: The Meanings Chilean Teachers Convey in the Neoliberal Context. *Sustainability*, 13(23), 13390. <https://doi.org/10.3390/su132313390>
- Rodrigo, L. (2019). Los programas internacionales de evaluación estandarizada y el tratamiento de sus datos a nivel nacional. El caso de Argentina en el estudio PISA de la OCDE. *Foro de Educación*, 17(26), 73-94. <https://doi.org/https://doi.org/10.14516/fde.708>
- Ruiz, I., y Petrova, D. (2019). Scoping reviews. Another way of literature review. *Medicina Clínica*, 153(4), 165-168. <https://doi.org/10.1016/j.medcle.2019.02.026>
- Sahlberg, P. (2016). The Global Educational Reform Movement and Its Impact on Schooling. En K. Mundy, A. Green, B. Lingard, y A. Verger, *Education and a Global Polity* (págs. 128-144). Wiley: Blackwell. <https://doi.org/10.1002/9781118468005.ch7>
- Salazar, J., y Rifo, M. (2020). Bajo el péndulo de la historia: Cinco décadas de masificación y reforma en la educación superior chilena (1967-2019). *Economía y Política*, 7(1), 27-61. <https://dialnet.unirioja.es/servlet/articulo?codigo=8737814>
- San Martín, V., Paukner, F., y Sanhueza, S. (2015). ¿Cuál cable primero? El desarme del sistema educativo en Chile. *Revista de Estudios Latinoamericanos*, 60, 159-185. <https://doi.org/10.1016/j.larev.2015.03.002>

- Strello, A., Strietholt, R., Steinmann, I., y Siepmann, C. (2021). Early tracking and different types of inequalities in achievement: difference-in-differences evidence from 20 years of large-scale assessments. *Educational Assessment, Evaluation and Accountability*, 33, 139-167. <https://doi.org/10.1007/s11092-020-09346-4>
- Telling, K. (2020). The complexity of educational elitism: moving beyond misrecognition. *British Journal of Sociology of Education*, 41(7), 927-941. <https://doi.org/10.1080/01425692.2020.1789847>
- Vázquez, M. (2015). La calidad de la educación. Reformas educativas y control social en América Latina. *Revista de Estudios Latinoamericanos*, (60), 93-124. <https://doi.org/10.1016/j.larev.2014.10.001>
- Westphaln, K., Regoeczi, W., Masotya, M., Vazquez, B., Lounsbury, K., McDavid, L., Lee, H., Johnson, J., y Ronis, S. (2021). From Arksey and O'Malley and Beyond: Customizations to enhance a team-based, mixed approach to scoping review methodology. *MethodsX*, 8. <https://doi.org/10.1016/j.mex.2021.101375>
- Williamson, B. (2021). Making markets through digital platforms: Pearson, edu-business, and the (e)valuation of higher education. *Critical Studies in Education*, 62(1), 50-66. <https://doi.org/10.1080/17508487.2020.1737556>
- Williamson, B., Bayne, S., y Shay, S. (2020). The datafication of teaching in Higher Education: critical issues and perspectives. *Teaching in Higher Education*, 25(4), 351-365. <https://doi.org/10.1080/13562517.2020.1748811>

FINANCIACIÓN

Los autores no recibieron financiación para el desarrollo de la presente investigación.

DECLARACIÓN DE CONFLICTO DE INTERÉS

Ninguno.

AGRADECIMIENTOS

Marcela Muñoz Lira agradece el apoyo de Conicyt PFCHA/Magister Profesionales de la Educación Nacional 50190039.

CONTRIBUCIÓN DE AUTORÍA

Conceptualización: Marcela Muñoz Lira y Julio Bruna Gaete.

Curación de datos: Marcela Muñoz Lira y Julio Bruna Gaete.

Análisis formal: Marcela Muñoz Lira y Julio Bruna Gaete.

Adquisición de fondos: Marcela Muñoz Lira y Julio Bruna Gaete.

Investigación: Marcela Muñoz Lira y Julio Bruna Gaete.

Metodología: Marcela Muñoz Lira y Julio Bruna Gaete.

Administración del proyecto: Marcela Muñoz Lira y Julio Bruna Gaete.

Recursos: Marcela Muñoz Lira y Julio Bruna Gaete.

Software: Marcela Muñoz Lira y Julio Bruna Gaete.

Supervisión: Dr. Félix Ángulo Rasco, profesor del curso Evaluación y Estándares del magister en Educación, con mención en Evaluación Educativa, PUCV.

Validación: Marcela Muñoz Lira y Julio Bruna Gaete.

Visualización: Marcela Muñoz Lira y Julio Bruna Gaete.

Redacción – borrador original: Marcela Muñoz Lira.

Redacción – revisión y edición: Marcela Muñoz Lira y Julio Bruna Gaete.