

LA EVALUACIÓN EN LA MIRA

Una examinación crítica y propositiva para la
evaluación nacional, docente y de aula

LA EVALUACIÓN EN LA MIRA

Una examinación crítica y propositiva para la evaluación nacional, docente y de aula

Alejandra Falabella, María Beatriz Fernández y María Teresa Flórez Petour
Editoras

Ediciones Universidad Alberto Hurtado
Alameda 1869 · Santiago de Chile
mgarciam@uahurtado.cl · 56-228897726
www.uahurtado.cl

Marzo 2025

Los libros de Ediciones UAH poseen tres instancias de evaluación: comité científico de la colección, comité editorial multidisciplinario y sistema de referato ciego. Este libro fue sometido a las tres instancias de evaluación.

Este libro fue financiado gracias al aporte de la Facultad de Educación de la Universidad Alberto Hurtado y ANID/PIA/Fondos Basales para Centros de Excelencia FB0003 de la Universidad de Chile.

ISBN libro impreso: 978-956-357-510-1

ISBN libro digital: 978-956-357-511-8

Coordinadora Colección Educación
María Teresa Rojas

Dirección editorial
Alejandra Stevenson Valdés

Editora ejecutiva
Beatriz García-Huidobro

Diseño interior y portada
Alejandra Norambuena

Portada
iStock



Grupo de
Editoriales
Universitarias
AUSJAL

Con las debidas licencias. Todos los derechos reservados. Bajo las sanciones establecidas en las leyes, queda rigurosamente prohibida, sin autorización escrita de los titulares del copyright, la reproducción total o parcial de esta obra por cualquier medio o procedimiento, comprendidos la reprografía y el tratamiento informático, así como la distribución de ejemplares mediante alquiler o préstamos públicos.

LA EVALUACIÓN EN LA MIRA

Una examinación crítica y propositiva para la
evaluación nacional, docente y de aula

ALEJANDRA FALABELLA
MARÍA BEATRIZ FERNÁNDEZ
MARÍA TERESA FLÓREZ PETOUR
EDITORAS

ÍNDICE

INTRODUCCIÓN

La evaluación exacerbada

*Alejandra Falabella, María Beatriz Fernández
y María Teresa Flórez Petour*

11

PRIMERA PARTE

EVALUACIONES NACIONALES E INTERNACIONALES

29

CAPÍTULO 1

La expansión global de las pruebas de desempeño: instrumentación y trayectorias en tiempos de rendición de cuentas y aseguramiento de la calidad educativa

Antoni Verger, Clara Fontdevila y Lluís Parcerisa

31

CAPÍTULO 2

Evaluaciones internacionales a gran escala en Chile: métricas globales y sus (des)contentos

Cristina Alarcón López

67

CAPÍTULO 3

El gobierno de las ilusiones: un análisis crítico del Simce

Paulina Contreras, Felipe Acuña y Alejandra Falabella

107

CAPÍTULO 4

Evaluando la evaluación: tensiones en el debate por la validez de las pruebas estandarizadas

María Teresa Flórez Petour

143

CAPÍTULO 5

¿Se puede? Modelos alternativos de evaluaciones a gran escala

Javier Campos Martínez y Fernanda Gándara

179

SEGUNDA PARTE

EVALUACIÓN DE LA FORMACIÓN Y EL EJERCICIO DOCENTE

211

CAPÍTULO 6

Uso de la rendición de cuentas como reforma: política y práctica de la formación del profesorado

Elizabeth S. Keefe y Andrew F. Miller

213

CAPÍTULO 7

Formación inicial docente para la justicia social: desafíos en tiempos de rendición de cuentas

*María Beatriz Fernández, Ilich Silva-Peña
y César Peña-Sandoval*

245

CAPÍTULO 8

La evaluación nacional diagnóstica (END) en la formación inicial docente desde la perspectiva de los evaluados: elementos para una revisión de la política

Andrea Ruffinelli, Tatiana Cisternas, Carla Förster y Francisca Donoso

271

CAPÍTULO 9

Contradicciones y paradojas de las trayectorias docentes: entre la evaluación con altas consecuencias y la cotidianidad escolar

*Claudia Carrasco-Aguilar, Lilian Vergara, Verónica López,
Tabisa Verdejo y Sebastián Ortiz*

303

CAPÍTULO 10

Consecuencias sociopsíquicas de la evaluación neoliberal-managerial

*Patricia Guerrero, Leandro Silva, Mauro Ramos,
Oriana Muñoz y Javiera Salinas*

327

TERCERA PARTE

EVALUACIÓN EN EL AULA

355

CAPÍTULO 11

Aproximaciones al campo de la investigación y las prácticas de evaluación en el aula

Rebeca Anijovich y Graciela Cappelletti

357

CAPÍTULO 12

La política evaluativa y sus consecuencias en el aula: dificultades y posibilidades para avanzar hacia una evaluación con foco formativo

Tamara Rozas y Elisa De Padua

391

CAPÍTULO 13

Experiencias de evaluación formativa en Chile: de un concepto aislado a un enfoque integral

José Miguel Olave, Gloria Contreras y Ximena Azúa

427

CAPÍTULO 14

Evaluación de aula para la justicia social en contextos indígenas: el caso de estudiantes mapuche

*Miguel Del Pino, Silvia Castillo, Maximiliano Heeren
y Pamela Tejeda*

451

CAPÍTULO 15

Evaluación inclusiva: ¿cómo dar respuesta a la diversidad en el aula para promover una educación con sentido?

Constanza San Martín, Rocío Rojas†, Liliana Ramos y Luz María Terán

481

A MODO DE CIERRE

Reimaginar un nuevo paradigma evaluativo

*Alejandra Falabella, María Beatriz Fernández
y María Teresa Flórez Petour*

511

INTRODUCCIÓN

LA EVALUACIÓN EXACERBADA

*Alejandra Falabella, María Beatriz Fernández
y María Teresa Flórez Petour*

Exámenes nacionales e internacionales a estudiantes de educación escolar, pruebas de conocimiento a docentes, portafolios docentes, grabaciones de aula, pruebas a estudiantes de pedagogía, acreditación a carreras universitarias de pedagogía, indicadores de desarrollo personal y social, visitas evaluativas a establecimientos, pruebas de ensayo y monitoreo, etc. Todo ello es parte de la exacerbación evaluativa del sistema educativo chileno, similar a lo que ocurre en distintos lugares del mundo.

Es una “vida datificada” en palabras de Deborah Lupton (2018), con masivos volúmenes de datos, puntajes, indicadores, “evidencias”, sobre lo que hacemos (y lo que no hacemos), cómo lo hacemos y con qué resultados. Estas son tecnologías políticas que no son inocuas, comúnmente asociadas a metas de trabajo, sistemas de monitoreo y comparación, y vinculadas a consecuencias, como bonos, símbolos de prestigio y desprestigio, amenazas y sanciones, como también asesorías técnicas y exigencias de formación continua.

Las y los docentes de escuela —seguramente más que nadie— entienden y experimentan esta “gobernanza de la métrica evaluativa” (Falabella, 2023). Las métricas sitúan la identidad docente en tensión entre los sentidos que otorgan a su labor y la necesidad de adaptar sus prácticas a las lógicas de la evaluación externa y su pedagogía subyacente. Es una tensión entre una pedagogía centrada en el proceso, y una centrada en los resultados; entre una

formación holística y una con foco en los contenidos a evaluar; y entre innovar nuevas metodologías y seguir pautas estandarizadas. Los efectos de estas políticas transforman no solo las prácticas pedagógicas, pero también los modos de entender los propósitos de la escuela, y las maneras de tomar decisiones en la gestión pedagógica.

Las y los estudiantes de escuela también saben de métricas. En el caso de Chile, desde primer año básico ya deben aprender a entender lo que es un 4,2 o un 5,6 o un 7,0, sumado a su involucramiento en las pruebas nacionales estandarizadas que pueden comenzar desde los 7 años de edad (segundo básico). Así, el interés inicial por el aprendizaje y el descubrimiento es desplazado paulatinamente hacia “el valor de cambio” de la calificación, según argumenta Santos Guerra (2003), con aulas que comienzan a ser habitadas por frases como: “¿profe, esto es con nota?”, “¿se darán décimas extras por esta actividad?”, “¿esto entra en la prueba?”. Cuando el foco de la evaluación está en esta sintonía, se promueven identidades reputacionales entre estudiantes (Stobart, 2010), quienes asocian la evaluación a características personales estáticas (“soy malo” o “bueno” para...). Se alimenta, de este modo, la pérdida de sentido de estar en el aula, afectando las identidades de los y las estudiantes como aprendices.

Por último, quienes somos académicas y académicos bien sabemos de esto; nos miden cuántas publicaciones tenemos al año, la indexación, el factor de impacto y el cuartil de las revistas en que publicamos (Fardella, 2020). Pero no solo eso, participamos del juego de la métrica. Nos ponemos en Academia.edu o en ResearchGate que, a su vez, de forma pública, como un escenario de espectáculo, nos miden, nos felicitan, nos ranquean y nos comparan. Sumergidos en la presión y emocionalidad del “ego métrico”, nos alegramos, frustramos y autoexplotamos. Los que, además, trabajamos en formación docente o en posgrados en universidades chilenas invertimos tiempo y energía para responder a la acreditación del Consejo Nacional de Acreditación (CNA), pues sabemos que tiene consecuencias en el prestigio de estos programas y para que sigan funcionando.

Esta gobernanza evaluativa cala hondo en los sistemas educativos, como tal vez en ningún otro servicio público. La educación escolar es, especialmente, un ámbito estratégico para los países y su futuro. A pesar de la propagación de las políticas de privatización, mercantilización y la disminución de la oferta pública, los Estados, en vez de disminuir su poder, han tendido a recrear sus mecanismos de control sobre el funcionamiento de la educación (Ball y Youdell, 2008; Maroy, 2009; Ozga, 2016). El poder estatal no disminuye, sino que cambia sus formatos y tecnologías. En este escenario, la evaluación es una de las herramientas políticas de regulación masiva más ampliamente extendidas sobre los sistemas educativos.

Las evaluaciones no son meramente “un instrumento” o “una medición”, delimitadas a su aplicación y resultados. Tampoco son una fase, como parte de un secuenciado y ordenado “círculo virtuoso”, en que los sujetos (casi vacíos de saber) se iluminan con los resultados, para luego disponerse a establecer estrategias de mejoramiento. La realidad, como se ilustra a lo largo de los capítulos de este libro, es mucho más compleja e intrincada.

La evaluación es poder. Establece horizontes de lo deseable, es capaz de girar prioridades institucionales y deseos personales. Es un poder performático sobre los actores del sistema. No es solamente un resultado, un número listo para ser leído y gatillar acciones “de mejora”. La gobernanza de la métrica funciona como un dispositivo de poder, en términos foucaultianos (Foucault 2006, 2008; Rose, 1996), induce a discursos, prácticas y procesos de subjetivación. La evaluación se convierte en una métrica que nos constituye.

La gubernamentalidad, explica Ramos (2019), es un conjunto de teorías, instituciones, procedimientos, análisis, cálculos, estrategias y tecnologías que permiten regular la conducta de poblaciones, posibilitando así la “conducción de las conductas”. Detrás de la gubernamentalidad de la métrica subyace una lógica capitalista de la tecnocracia managerial, de acumulación, comparación y competencia. Son “modos de acción sobre sí mismo” en palabras de Zangaro (2011), con un carácter eminentemente activo en la construcción de los sujetos manageriales. La eficacia en estos dispositivos está en un poder que individualiza a los sujetos por medio de

responsabilizar, culpar o enaltecer. Es un poder que captura la identidad de las personas, el cual es difícil de eludir (Falabella, 2014).

El poder de los números —y de las clasificaciones— es que se naturalizan. La métrica se hace invisible y parece autoevidente, se construye el sentido común de “lo bueno”, “lo deseable” y “lo despreciable”. Es “el espíritu de la cuantificación” explica Power (2004). Esta narrativa procura orden, exactitud y certidumbre en un campo, justamente, complejo, incierto e inexacto como son los procesos pedagógicos. Es una ecuación simple, económica y prometedora. Ello es parte de la “numerolatría”, de la seductora fantasía evaluativa de simplificar y reducir la realidad.

Es un conocimiento monolítico que crea la “tiranía de los números” (Ball, 2015) como criterio de decisión. Confunde lo político y lo técnico, y se pierden los sentidos democráticos de la evaluación. De allí la profundidad y complejidad en estudiar las evaluaciones. Aunque, a su vez, los actores tienen agencia y son creativos, interpretan, esconden, destacan, cuestionan y discuten las métricas (Campos-Martínez y Fernández, 2014; Villalba *et al.*, 2023), como se evidencia en los capítulos de este libro.

Este contexto exige hacer una pausa y discutir las preguntas de fondo: ¿el intenso entramado evaluativo mejora los procesos educativos? Tanto tiempo, recursos y energía invertida en evaluar en distintos niveles —global, nacional, aula—, y a distintos actores; ¿tiene sentido?, ¿contribuye a una mejor educación?, ¿tiene consecuencias deseables? y ¿qué repertorios alternativos emergentes existen en la experiencia nacional e internacional?

Este es el propósito de este libro. Poner la evaluación en la mira, bajo una lupa, examinarla críticamente en detalle, a partir del conocimiento acumulado en el área y, a su vez, avanzar en proponer nuevos escenarios evaluativos posibles.

EL CASO DE CHILE

Las políticas de evaluación estandarizada a gran escala han viajado y se han expandido de forma masiva en los sistemas educativos a

nivel global (Verger *et al.*, 2019, capítulo 1 de este libro). América Latina tuvo una extendida agenda de promoción de los sistemas de evaluación a gran escala, bajo el lema de la “calidad y la equidad”, durante los años de 1990 y 2000, secundada por las organizaciones internacionales. En este libro profundizamos en el caso de Chile, como un caso emblemático, en la región y en el mundo, de un sistema exacerbado de evaluaciones con altas consecuencias.

Chile tiene una larga historia de evaluaciones en educación (Falabella y Ramos, 2019; Flórez, 2014). Desde los inicios del sistema educativo se comienza lentamente, desde la segunda mitad del siglo XIX, a crear mecanismos de evaluación estatal, como el sistema de inspección a las escuelas primarias (Soifer, 2009). Los llamados visitadores viajaban a través de las provincias para inspeccionar las escuelas, revisaban la higiene de los niños y niñas, evaluaban los avances de los alumnos y los métodos de enseñanza del profesorado, y examinaban los libros e información estadística de la escuela. Todo ello con un velo de solemnidad, pues representaban nada menos que el naciente Estado republicano en las precarias escuelas de la época. Estas inspecciones estaban vinculadas a consecuencias que incluían la destitución del cargo de un docente si era mal evaluado (Egaña, 2000; Falabella y Ramos, 2019). En definitiva, el sistema educativo tempranamente contaba con un sistema de evaluación centralizado, con altas consecuencias sobre los establecimientos educativos y la labor de los y las docentes primarios.

Por otra parte, se crearon evaluaciones externas para el estudiantado, como mecanismos de control y certificación. A nivel de educación primaria, al final del ciclo, se llevaban a cabo exámenes orales por asignatura, ante comisiones designadas por el gobernador. Paralelamente, los liceos de la educación secundaria —públicos y privados— estaban supervigilados por la Universidad de Chile, y controlados por los “exámenes finales”, que luego se transformarían en el “Bachillerato” para el ingreso a la universidad.

En cuanto al sistema de calificaciones, en el siglo XIX se utilizaban las categorías de distinción, aprobado o reprobado como las primeras etiquetas sobre el rendimiento de los y las estudiantes, según el criterio de los examinadores externos. Luego, el Decreto

Disposiciones Relativas al Servicio de Instrucción Primaria de 1899 estableció que los examinadores asignaban notas del 1 al 4, siendo el 1 la nota más alta, equivalente a un nivel muy bueno, y el 4 la más baja, equivalente a malo. El Decreto N° 2.545 de 1929, posteriormente, reglamentó la escala numérica del 1 a 7 que conocemos actualmente, con variaciones en el tiempo acerca del punto en que se situaba el nivel de aprobación (Gysling, 2015). De este modo, los orígenes de la evaluación de los aprendizajes estaban marcados por la calificación para cumplir una función social y de certificación, y poco o nada se aludía a consideraciones pedagógicas y formativas.

Mientras tanto, el conocimiento de la psicometría y de los “exámenes de inteligencia” llegó al país a fines del siglo XIX e impactó en la educación desde inicios del siglo XX con la construcción de pruebas estandarizadas (Mayorga, 2018). Estas influencias hicieron sinergia con la pedagogía experimental, la educación progresiva y una racionalidad científica y laica del Estado Docente, asociada a este tipo de evaluaciones entendidas como “científicas”. Tanto así que en 1930, explica Irma Salas (1942), se aplicó la primera prueba experimental a gran escala en la educación secundaria desde el Ministerio de Educación Pública (Falabella y Ramos, 2019).

Las evaluaciones descritas, en los distintos niveles del sistema, serían los inicios de la gobernanza de la métrica en el país. Ahora bien, ya en esos años, las publicaciones en revistas, desarrolladas por docentes, levantaban las tensiones que estos dispositivos generaban. Se señalaba que, si bien los exámenes tenían un potencial de “grandes ventajas para la escuela”, conllevaban la posibilidad de “poner un estigma al niño, un sello de inferioridad” (Velásquez, 1927a, pp. 70-71), en el marco de lo que llamaban una posible “aristocracia de la inteligencia”, que terminaría fomentando la exclusión (Velásquez, 1927b, pp. 97-98).

En la década de los 60, con la masificación de la educación y la planificación estatal, emerge otra época efervescente en cuanto a las evaluaciones nacionales y a nivel de aula. En este tiempo se pasa del Bachillerato a la Prueba de Aptitud Académica, la primera prueba nacional estandarizada para el ingreso a la educación

superior en el año 1967. Junto a ello, a fines del mismo año, se experimenta con la aplicación de la Prueba Nacional de Habilidad Verbal y Matemática como mecanismo para certificar el cumplimiento de habilidades para el egreso del octavo año básico. Sin embargo, en la práctica aquello fue un propósito fallido, según explicó Mario Leyton, quien lideró el diseño de la prueba (Leyton, 1970). Ambos exámenes, de todas formas, tenían el fin de generar consecuencias individuales en las y los estudiantes, acorde a sus resultados. Conjuntamente, el año 1971, se aplicó la primera prueba internacional —*Six Subject Study*— dirigida por la *International Association for the Evaluation of Educational Achievement* (IEA).

A nivel de aula, se introdujo el uso de “evaluaciones objetivas”, es decir, pruebas con ítems de respuestas cerradas o estructuradas (por ejemplo: sección múltiple, verdadero o falso, términos pareados). Ello significaba una revolución pedagógica, influenciada por el currículum tecnológico, en que se dejaba en un segundo plano la larga tradición de exámenes orales y de ensayos, interpretadas desde aquel lente teórico como evaluaciones subjetivas y precientíficas (Flórez, 2014).

Lo anterior se cruza con la entrada de la evaluación como tema específico en la formación inicial docente en Chile, cuyo comienzo estaba más relacionado con elementos de medición y estadística aplicados a la educación. La formación de docentes de educación secundaria en estas materias tiene sus primeras apariciones en los años 30, mientras que para las y los docentes de educación primaria circulaba más bien en el formato de materiales tipo manuales (Gysling, 2015). Así, en diálogo con los desarrollos teóricos e investigativos a nivel internacional, la noción de evaluación formativa, asociada a la función pedagógica de la evaluación de aula, se instala recién a partir de los años 60 y 70, desde las primeras nociones neoconductistas del concepto a partir de los planteamientos de Bloom, Hastings y Madaus (1977). Por medio de manuales y cursos, los y las docentes en Chile se acercaron a una noción de evaluación formativa, en una lógica de corrección de errores y mecanismos de reforzamiento en torno a criterios explícitos (Flórez, 2014).

Bajo la dictadura cívico-militar y su política neoliberal durante los años 80, se continuó con el desarrollo de dispositivos evaluativos a gran escala: primero el Programa de Evaluación del Rendimiento Escolar —PER— (1982) y luego el Sistema de Medición de la Calidad de la Educación —Simce— (1988). El giro político con la creación del Simce no era solamente que fuese para proveer información pedagógica o controlar la aplicación del currículum nacional en las escuelas, sino que, bajo la racionalidad neoliberal de la época, se diseñó para alimentar el nuevo mercado educativo, con un financiamiento estatal competitivo. La expectativa era, y sigue siendo en parte, orientar la elección escolar (y salida) de las familias a partir de los resultados por establecimiento, y de este modo regular la calidad del sistema educativo por medio de la “mano invisible” de la (informada) demanda (Campos *et al.*, 2015; Falabella, 2021; Prieto, 1983).

Gradual y especialmente a partir de la década de los 2000, la lógica de la gobernanza de la métrica y la intensidad evaluativa se impregnó y aumentó en casi todos los rincones del sistema escolar. Esto consolidó lo que se denomina un Estado evaluador o hipervigilante (Bellei, 2015; Maroy, 2009; Falabella, 2014, 2021). Como resultado, contamos con una larga lista de evaluaciones, cada una asociada a redes de mediciones, categorizaciones, sistemas de prestigio y desprestigio, y premios, amenazas y sanciones. Es un “mercado intensificado”, en que el sistema de evaluación constituye un modo de ordenamiento social, según argumenta Santori (2023). A continuación, intentamos resumir este entramado de evaluaciones presentes en el sistema educativo escolar actual.

Partimos primero por la prueba Simce, que es un eje troncal de la gobernanza de la métrica del sistema escolar. Este se aplica en alrededor de ocho a diez pruebas anuales censales para estudiantes desde segundo año básico a segundo medio. Los “Planes de Mejoramiento”, la evaluación a directivos, la clasificación de los establecimientos según desempeño, los bonos docentes del Sistema Nacional de Evaluación de Desempeño (SNED), la Subvención Escolar Preferencial (SEP), entre otros, están engarzados a la métrica del Simce. Asimismo, los resultados por establecimiento son

publicados en la página web de la Agencia de Calidad de la Educación¹ y, por medio de la Ley N° 20.529 del Sistema de Aseguramiento de Calidad (2011), se define que los establecimientos con bajos resultados reiterados recibirán consecuencias, lo que incluye el cierre del establecimiento.

Sumado a ello, contamos con los Indicadores de Desarrollo Personal y Social (nombrados por ley como los “Otros indicadores de la calidad”) que tienen el propósito de evaluar aspectos tales como autoestima académica, retención escolar, formación ciudadana, brechas de género, entre otros, que se basan en indicadores de los establecimientos y en una encuesta respondida por apoderados y estudiantes. Estos indicadores inciden en la categorización de los establecimientos en el sistema de ordenamiento según desempeño, pero su peso en comparación con los indicadores derivados del Simce en el proceso de ordenación es significativamente menor². Además, existen las “visitas evaluativas” llevadas a cabo por profesionales de la Agencia de la Calidad, con el propósito de evaluar in situ la calidad del desempeño de los establecimientos escolares con una baja categorización según desempeño.

Otra evaluación que se suma a este listado es el Diagnóstico Integral de Aprendizajes (DIA), autoaplicada tres veces al año en los establecimientos, con el fin de monitorear los aprendizajes de los y las estudiantes y ofrecer una retroalimentación pedagógica. Además, están las pruebas internacionales, por ejemplo, el Programa Internacional de Evaluación de Estudiantes (PISA) de la OCDE y el Estudio de las Tendencias en Matemáticas y Ciencias (Timss) de la Asociación Internacional para la Evaluación del Rendimiento Educativo, como también la Prueba de Acceso a la Educación Superior (PAES), que también tensiona la enseñanza de la educación media (Gazmuri *et al.*, 2022).

¹ Institución a cargo de evaluar y orientar el mejoramiento de los establecimientos de educación parvularia y escolar.

² La Ley N° 20.529 estipula que el puntaje Simce pondera un 67% versus un 33% los Indicadores de Desarrollo Personal y Social para la ordenación del desempeño de los establecimientos escolares. Más aún, debe considerarse que estos últimos indicadores también incluyen el incremento de resultados Simce, por lo que en efecto el Simce pesa un 73% versus un 27% (Oyarzún y Falabella, 2022).

En cuanto a los y las docentes, se agrega otro paquete de evaluaciones. Este listado comienza desde que son estudiantes de primer año en carreras de pedagogía, siendo los únicos estudiantes universitarios que por ley deben rendir ya en primer año una prueba diagnóstica inicial³ diseñada por cada institución, y deben participar de acciones remediales si sus resultados no son satisfactorios. Además, les corresponde rendir en el penúltimo año de la carrera la Evaluación Nacional Diagnóstica, que incluye pruebas nacionales de conocimiento pedagógico, disciplinario y didáctico (Fernández y Madrid, 2020). Al mismo tiempo, las carreras de pedagogía son evaluadas por la Comisión Nacional de Acreditación⁴ (CNA) y deben ser acreditadas⁵ y estar alineadas con los estándares nacionales para la formación docente, que complementan a los estándares de acreditación, de lo contrario arriesgan ser cerradas, pues como consecuencia no se les permite matricular nuevos estudiantes.

Las y los docentes en ejercicio, del sector público y privado subvencionado, son evaluados por medio de la Carrera Docente (Ley N° 20.903, 2016), con pruebas de conocimiento disciplinar y pedagógico y un portafolio de su trabajo pedagógico. A esto, hasta el año 2023⁶, se agregaba la evaluación docente (Ley N° 19.961, 2004) del sector público, que constaba de cinco instrumentos que, sumando a los instrumentos evaluativos de la Carrera Docente, incluían una entrevista de evaluador par, la evaluación de un directivo del establecimiento y una autoevaluación. A pesar de la reducción de instrumentos, con la aprobación de un sistema único de evaluación docente (Ley N° 21.625, 2023), las consecuencias persisten. El sistema evaluativo de las y los docentes sigue estando

³ En la Ley N° 20.903 del Sistema de Desarrollo Profesional Docente se indica que se debe realizar una evaluación de las “creencias docentes” en primer año de pedagogía, aunque en la práctica estas pruebas son muy heterogéneas y su contenido depende de cada institución.

⁴ A cargo de evaluar y acreditar la calidad de las instituciones en la educación superior.

⁵ Este requisito existe para algunas carreras del área de la salud, hasta hoy en día en medicina y odontología.

⁶ En octubre de 2023 se promulga la Ley N° 21625 que establece un sistema único de evaluación docente.

vinculado a consecuencias, tanto al aumento de sueldo y entrega de bonos, como a sistemas de capacitación continua y potenciales despidos en caso de evaluaciones deficientes reiteradas.

En cuanto a la evaluación de aula, se mantiene hasta hoy la escala de notas (1 al 7), que se debe utilizar desde el primer año básico, reflejo de una larga tradición de la evaluación sumativa en el país. Los reglamentos y decretos actuales han tendido a centrarse en la regulación de las cantidades de calificaciones, la función sumativa de la evaluación y aspectos de la promoción o repitencia escolar. A ello se agrega, una serie de diagnósticos, evaluaciones y monitoreos para estudiantes con necesidades educativas especiales, definidos en el Decreto N° 170 del Programa de Integración Escolar (2009).

Ello se combina, al mismo tiempo, con el Decreto N° 67 (del año 2018) que es el primero en la historia evaluativa del país que escapa del foco de la calificación y promoción de los y las estudiantes para centrarse en los principios y componentes de una evaluación para el aprendizaje. El decreto modifica la tendencia hacia la repitencia automática, pasando a un modelo de repitencia excepcional, es decir, solamente se admite si hay un proceso de decisión colectiva de los actores de la escuela, donde se concluya que esta decisión va en beneficio del o de la estudiante.

Este enfoque evaluativo tuvo desarrollos teóricos iniciales en Chile, desde la década de los 80 hasta las actuales aproximaciones de base constructivista social, con enfoques como la evaluación auténtica, la evaluación como autorregulación y la evaluación para el aprendizaje. Pese a estas influencias y a la evidencia largamente acumulada acerca de sus efectos positivos (Flórez y Sammons, 2013), se trata de enfoques que no se han instalado con fuerza en la política educacional (Azúa Ríos, 2012). Asimismo, hasta el presente, la evaluación formativa experimenta dificultades para llegar al aula, tanto por los vacíos en la formación de los y las docentes, como por las tensiones que pone la propia política evaluativa en términos de mensajes divergentes y contradictorios acerca de cuál debe ser el foco del trabajo docente en este ámbito (Flórez, 2017).

SOBRE ESTE LIBRO

La métrica nos sumerge en la vida cotidiana del sistema educativo. A pesar de los cuestionamientos a este tipo de evaluaciones en distintos momentos de la historia en Chile, pareciera que vivimos una desmemoria de los debates de fondo y un retroceso en la posibilidad de reimaginar un nuevo paradigma evaluativo. Resulta frecuente escuchar declaraciones respecto a la “importancia de estar informados”, que “no podemos andar a ciegas”, “debemos tomar decisiones basadas en información objetiva”, “debemos ser responsables”. Son discursos que reducen la comprensión de la evaluación desde una racionalidad instrumental y exacerbaban la métrica, de un modo casi religioso. Bajo estos discursos se esconde la producción de las evaluaciones, categorizaciones y mediciones, las discusiones y dudas al construirlas, y se invisibiliza lo que no es posible de cuantificar, mientras el estatus de objetividad de los números resulta incuestionado.

Decidimos hacer este libro con el fin de cuestionar la lógica evaluativa imperante y contribuir a una evaluación que sea un aporte a los procesos de mejoramiento educativo y a la formación integral de los estudiantes. No será posible transformar la educación acorde a los desafíos del siglo XXI, sino transformamos el sistema evaluativo. A través de análisis basados en la investigación acumulada, examinamos tres ámbitos: i) evaluaciones de aprendizajes a gran escala internacionales y nacionales, ii) evaluaciones a la formación docente y docentes en ejercicio, y iii) la evaluación a nivel de aula. En cada una de estas tres secciones del libro se revisan críticamente los procesos de construcción de estas políticas y se muestra de qué modos se experimentan, se ponen en práctica e inciden estas evaluaciones en el trabajo pedagógico de la vida cotidiana de las instituciones educativas. Igualmente, las autoras y autores de los capítulos desarrollan perspectivas de futuro, en búsqueda de reimaginar otros modos de evaluar en educación.

La primera sección aborda las evaluaciones a gran escala. El capítulo 1, desde una perspectiva internacional, revisa las principales tendencias respecto a las políticas de evaluación y rendición de

cuentas, e identifica trayectorias diferenciadas entre países que han implementado estas medidas. El capítulo 2 examina el uso y los debates sobre las evaluaciones internacionales en Chile, desde sus orígenes hasta hoy en día. Los capítulos 3 y 4 realizan un análisis crítico respecto a la prueba Simce. El capítulo 3 examina sus efectos y transformaciones en la vida cotidiana de la escuela, y el capítulo 4 interroga la concepción de validez subyacente a la evaluación Simce, y propone una comprensión compleja y holística de validez, con el fin de repensar una evaluación a gran escala. El capítulo 5 realiza una revisión de sistemas evaluativos a gran escala alternativos al modelo de rendición de cuentas con altas consecuencias y da luces respecto a cómo avanzar hacia un modelo con foco en la profesionalización docente y la evaluación para el aprendizaje.

La sección dos, sobre evaluaciones para la formación docente y para docentes en ejercicio, comienza con el capítulo 6, que problematiza las políticas de rendición de cuentas en la formación y el trabajo docente como el principal instrumento de reforma a nivel internacional, tomando como ejemplo el caso de Estados Unidos. Los capítulos siguientes abordan las evaluaciones nacionales más relevantes sobre la formación y el trabajo docente en Chile. Los capítulos 7 y 8 profundizan en las principales evaluaciones usadas para la rendición de cuentas en la formación docente, las acreditaciones y evaluaciones diagnósticas a los y las estudiantes de pedagogía, discuten sus implicancias para proyectos formativos que promueven la justicia social y sus principales desafíos desde la perspectiva de los y las estudiantes. Los capítulos 9 y 10 profundizan en la evaluación de los y las docentes como una evaluación de altas consecuencias y analizan sus implicancias para las trayectorias de desarrollo profesional docente y para los aspectos sociopsíquicos de los y las profesoras.

La tercera sección del libro se centra en la evaluación a nivel de aula, en el marco de las tensiones para la promoción de un enfoque formativo en evaluación, cuyos principios y principales problemáticas se presentan en el capítulo 11, desde una perspectiva internacional. Luego, el capítulo 12 sitúa estas tensiones en el contexto chileno, específicamente entre la función pedagógica

y la función de certificación y calificación, y entre la evaluación de aula y a gran escala. El capítulo 13 introduce formas de pensar en cómo transformar las culturas evaluativas de las comunidades escolares en Chile, a partir de la revisión de diferentes experiencias de desarrollo profesional docente, desde las cuales es posible tomar lecciones para promover cambios en un sentido formativo. Finalmente, los capítulos 14 y 15 ofrecen caminos y abren horizontes para un nudo crítico altamente presente en las aulas chilenas, y para los cuales las respuestas en evaluación no parecen suficientes: evaluación y diversidad. El capítulo 14 lo hace en un sentido más amplio, desde una propuesta para el abordaje inclusivo de la evaluación en el aula, mientras que el capítulo 15 se enfoca en las claves de una evaluación con perspectiva indígena para pensar en una evaluación con mayor justicia social.

REFERENCIAS BIBLIOGRÁFICAS

- Azúa Ríos, X. (2012). *¿Qué se evalúa cuando se evalúa? Una experiencia de formación docente en evaluación para el aprendizaje*. Ed. Universitaria.
- Ball, S.J. (2015). Education, governance and the tyranny of numbers. *Journal of Education Policy*, 30(3), 299-301.
- Ball, S. J. y Youdell, D. (2008). *Hidden privatization in public education*. Education International.
- Bellei, C. (2015). *El gran experimento: Mercado y privatización de la educación chilena*. Lom.
- Bloom, B., Hastings, T. y Madaus, G. (1977). *Evaluación del aprendizaje*. Troquel.
- Campos, J., Corbalán, F. y Inzunza, J. (2015). Mapping neoliberal reform in Chile: following the development and legitimation of the Chilean system of school quality measurement. En W. Au y J. Ferrare, *Mapping Corporate Education Reform* (pp. 106-125). Routledge.
- Campos-Martínez, J. y Fernández, B. (2014). Resistiendo la emboscada neoliberal desde las escuelas: Experiencias de organización docente contra las pruebas estandarizadas. *Revista Docencia*, 52(1), 72-82.
- Congreso Nacional de Chile (2004). *Ley N° 19.961. Evaluación Docente*. Biblioteca del Congreso Nacional.

- Congreso Nacional de Chile (2009). Decreto N° 170. *Normas para determinar los alumnos con necesidades educativas especiales que serán beneficiarios de las subvenciones para educación especial*. Biblioteca del Congreso Nacional.
- Congreso Nacional de Chile (2011). *Ley N° 20.529. Sistema Nacional de Aseguramiento de la Calidad de Educación Parvularia, Básica y Media y su fiscalización*. Biblioteca del Congreso Nacional.
- Congreso Nacional de Chile (2016). *Ley N°20.903. Crea el Sistema de Desarrollo Profesional Docente y modifica otras normas*. Biblioteca del Congreso Nacional.
- Congreso Nacional de Chile (2018). *Decreto 67, Aprueba normas mínimas nacionales sobre evaluación, calificación y promoción*. Biblioteca del Congreso Nacional.
- Congreso Nacional de Chile (2023). *Ley N° 21.625. Sistema único de evaluación docente*. Biblioteca del Congreso Nacional.
- Egaña, M^a L. (2000). *La educación primaria popular en el siglo XIX en Chile: Una práctica de política estatal*. Lom.
- Falabella, A. (2014). La escuela performativa: los efectos de las políticas de mercado y de responsabilización con altas consecuencias. *Education Policy Analysis Archives*, 22(51). <https://goo.su/LSRB9>.
- Falabella, A. (2021). The Seduction of *Hyper-Surveillance*: Standards, Testing, and Accountability. *Educational Administration Quarterly*, 57(1), 113-142.
- Falabella, A. (2023). El poder de los números en educación: subjetivización, resistencia y nuevos horizontes. En M.^a V. Leiva Guerrero y M. Pasqual Schneider (Comps.), *Políticas de evaluación y accountability en América Latina* (pp. 59-71). Ediciones Altazor.
- Falabella, A. y Ramos, C. (2019). La larga historia de las evaluaciones nacionales a nivel escolar en Chile. *Cuadernos Chilenos de Historia de la Educación*, 11, 66-98.
- Fardella, C. (2020). Abrir la jaula de oro. La universidad managerial y sus sujetos. *Izquierdas* 49, 2299-2320.
- Fernández, M. B, y Madrid, R. (2020). Profesionalización desde una perspectiva democrática: nudos críticos y propuestas para formación y desarrollo docente. En M.T. Corvera y G. Muñoz (Eds.), *Horizontes y propuestas para transformar el sistema educativo chileno* (pp. 206-233). Biblioteca del Congreso Nacional.

- Flórez, M.^a T. 2014. *Assessment reform in Chile: A contested discursive space*. Tesis Doctoral. Oxford University.
- Flórez, M.^a T. (2017). High-stakes assessment systems as a historical barrier in the struggle for change in education: the case of Chile. En M. Lawn y C. Alarcón (Eds.), *Cultures of Student Assessment* (pp. 213-234). Editorial Peter Lang.
- Flórez, M.^a T. y Sammons, P (2013). *A literature review of Assessment for Learning: effects and impact*. University of Oxford, Department of Education and CfBT.
- Foucault, M. (2006). Ethics: Subjectivity and truth. En P. Rabinow (Ed.), *Essential works of Foucault, 1954-1984* (Vol. 1). New Press.
- Foucault, M. (2008). *Birth of biopolitics (Michel Foucault: Lectures at the College De France)*. Palgrave Macmillan.
- Gazmuri, R., Toledo, M. I., Muñoz, E., y Ocampo, S. (2022). Las paradojas de la Prueba de Selección Universitaria de Historia y Ciencias Sociales: Dilemas y decisiones curriculares y pedagógicas de profesores chilenos. *Archivos Analíticos de Políticas Educativas*, 30 (124).
- Gysling, J. (2015). The historical development of educational assessment in Chile: 1810-2014. *Assessment in Education: Principles, Policy & Practice*, 23(1), 8-25.
- Leyton, M. (1970). La Experiencia Chilena. La Reforma Educacional: 1965-1970. Vol 1. CPEIP.
- Lupton, D. (2018). How Do Data Come to Matter? Living and Becoming with Personal Data. *Big Data and Society*, 5(2), 1-11.
- Maroy, C. (2009). Convergences and hybridization of educational policies around “post-bureaucratic” models of regulation. *Compare: A Journal of Comparative and International Education*, 39(1), 71-84.
- Mayorga, R. (2018). Las grandes reformas pedagógicas. En S. Serrano, M. Ponce de León, F. Rengifo y R. Mayorga (Eds.), *Historia de la educación en Chile (1810-2010). Tomo III. Democracia, exclusión y crisis (1930-1964)* (pp. 209-254). Taurus.
- Oyarzún, G., y Falabella, A. (2022). Indicadores de Desarrollo Personal y Social: La ilusión de la evaluación integral de la calidad. *Psicoperspectivas*, 21(1), 1-14.
- Ozga, J. (2016). Trust in Numbers? Digital Education Governance and the Inspection Process. *European Educational Research Journal*, 15(1), 69-81.

- Power, M. (2004). Counting, control and calculation: Reflections on measuring and management. *Human Relations*, 57(6), 765-783.
- Prieto, A. (1983). *La modernización educacional*. Ediciones Universidad Católica de Chile.
- Ramos, C. (2018). Dispositivo de evaluación y gubernamentalidad del sistema educacional: entretejimiento de ciencia social y poder. *Revista Cinta de Moebio*, 61, 41-55.
- Rose, N. (1996). Governing “advanced” liberal democracies. En A. Barry, T. Osborne, y N. Rose (Eds.), *Foucault and political reason: Liberalism, neoliberalism and rationalities of government* (pp. 37-64). UCL Press.
- Salas, I. (1942). *La investigación pedagógica en Chile*. Octavo Congreso Científico Americano (Washington, mayo de 1940). Anales de la Universidad de Chile, 45/46.
- Santori, D. (2023). *The quantified school: pedagogy, subjectivity and metrics*. Palgrave Macmillan.
- Santos Guerra, M.A. (2003). Dime cómo evalúas y te diré qué tipo de profesional y de persona eres. *Revista Enfoque Educativos*, 5(1), 69-80.
- Soifer, H. (2009). The Sources of Infrastructural Power. Evidence from Nineteenth-Century Chilean Education. *Latin American Research Review*, 44(2), 158-180.
- Stobart, G. (2010). *Tiempos de pruebas: los usos y abusos de la evaluación*. Morata.
- Velázquez, C. (1927a). Los tests psicológicos y la nueva educación. *Nuevos Rumbos*, 4, 97-98, 69-71.
- Velásquez, C. (1927b). Controversia entre aristócratas y demócratas de la inteligencia. *Nuevos Rumbos*, 5, 97-98.
- Verger, A., Parcerisa, L., y Fontdevila, C. (2019). The growth and spread of large-scale assessments and test-based accountabilities: A political sociology of global education reforms. *Educational Review*, 71(1), 5-30.
- Villalba-Pino, D., Valdés-Flores, F., y Fernández, M. B. (2023). Desafiando las nociones de trabajo docente: Críticas, propuestas y bilingüismos desde la organización docente. *Estudios Pedagógicos*, 49(2), 373-392.
- Zangaro, M. (2011). Subjetividad y trabajo: el management como dispositivo de gobierno. *Trabajo y Sociedad*, XV(16), 163-177.