

Indicadores de Desarrollo Personal y Social: La ilusión de la evaluación integral de la calidad

Personal and Social Development Indicators: The illusion of assessing schools' quality

Gonzalo Oyarzún Vargas^{1*}, Alejandra Falabella²

1 Programa de Idiomas, Universidad Finis Terrae, Santiago, Chile

2 Facultad de Educación, Universidad Alberto Hurtado, Santiago, Chile

* goyarzun@uft.cl

Recibido: 05-enero-2021

Aceptado: 03-diciembre-2021

RESUMEN

Tras décadas de instauración progresiva de evaluaciones estandarizadas en el sistema escolar, el debate se ha polarizado por los cuestionamientos a sus efectos colaterales y efectividad para medir la “calidad educativa”. Como respuesta, en el último tiempo, se desarrolló un intento por construir evaluaciones que aporten a una visión integral de la calidad. En Chile, uno de los instrumentos claves para cumplir dicho anhelo es la medición de los “Indicadores de Desarrollo Personal y Social” (denominados por ley como “Otros Indicadores de Calidad Educativa”); sin embargo, existe escaso conocimiento respecto a esta evaluación anual. Este artículo indaga en la gestación y construcción de esta medición y analiza críticamente sus resultados y aporte para evaluar integralmente la calidad educativa. Con este propósito, se examinaron informes y documentos (163) legislativos, gubernamentales, de organismos internacionales y de prensa. El artículo devela las inconsistencias técnico-políticas de estos indicadores y problematiza sus limitaciones para conseguir una “evaluación integral” dentro de un modelo de rendición de cuentas con altas consecuencias.

Palabras clave: calidad de la educación, Chile, evaluación escolar, evaluación integral, indicadores de desarrollo personal y social, rendición de cuentas

ABSTRACT

After decades of a progressive installation of standardized testing in the educational system, the debate has polarized due to questionings regarding its side effects and effectiveness to measure the ‘educational quality’. Thus, a search has begun to create instruments that contribute to acquire a comprehensive evaluative vision during the last years. This desire is aimed to be materialized in Chile by creating the Individual and Social Development Indicators (named in the law as “Other Education Quality Indicators”). Nonetheless, there is scarce knowledge about them. This paper delves into the antecedents that lead their creation and analyzes critically its results and effectiveness to measure comprehensively quality in education. To do so, the review of 163 documents and reports (legal, governmental, from international organizations, press) were considered. This article sheds light on these indicators’ political and technical inconsistencies and problematizes their limitations to achieve a ‘comprehensive assessment’ within a high-stakes accountability system.

Keywords: accountability, Chile, comprehensive assessment, personal and social development indicators, quality in education, school evaluation

Financiamiento: FONDECYT 1170477: La gubernamentalidad del sistema educacional chileno y su dispositivo evaluador: redes de producción, decisiones controversiales y efectos performativos (1968-2019).

Citación recomendada: Oyarzún Vargas, G., & Falabella, A. (2022). Indicadores de Desarrollo Personal y Social: La ilusión de la evaluación integral de la calidad. *Psicoperspectivas*, 21(1).

<https://dx.doi.org/10.5027/psicoperspectivas-vol21-issue1-fulltext-2194>



Publicado bajo licencia [Creative Commons Attribution International 4.0 License](https://creativecommons.org/licenses/by/4.0/)

Durante las últimas décadas, las políticas de rendición de cuentas en materia de educación escolar se han extendido en distintos lugares del mundo (Verger et al., 2019), bajo la premisa de que es posible medir y mejorar el rendimiento académico y la calidad educativa mediante evaluaciones estandarizadas y un conjunto de incentivos y apoyos para los equipos educativos, sumado a potenciales sanciones en caso de obtener bajos resultados. La proliferación de este tipo de gobernanza, sin embargo, ha suscitado álgidos debates y críticas. Existe un cúmulo significativo de investigaciones en las últimas décadas respecto a los efectos dañinos de estas medidas (Au, 2007; Falabella, 2014; 2020; Flórez Petour et al., 2018; Rosenkvist, 2010; Verger & Parcerisa, 2017). Las principales consecuencias identificadas son: la reducción curricular, la estandarización de los procesos de enseñanza y aprendizaje, la segregación de los estudiantes según su desempeño académico, la estigmatización hacia los establecimientos que atienden sectores marginales, y el estrés y perjuicio en el bienestar de las comunidades educativas.

La investigación acumulada muestra escasa y contradictoria evidencia respecto a si efectivamente este tipo de medidas incrementa el aprendizaje en los ámbitos evaluados. A la vez, han surgido movimientos antiestandarización, liderados por organizaciones docentes y de madres y padres, contra el uso de pruebas masivas a estudiantes (Hursh et al., 2020; Parcerisa & Villalobos, 2020). Por otra parte, organismos internacionales, como UNESCO (2017), que años atrás habían recomendado el uso de estos exámenes, han alertado igualmente sus efectos negativos y plantean la necesidad de reformular sus usos en el sistema escolar.

En respuesta a lo anterior, diversos países han intentado evitar o reducir estas consecuencias y han propuesto evaluar la calidad escolar de modo más holístico. En consecuencia, otros mecanismos de evaluación han emergido; a saber: ampliar los ámbitos de evaluación en las pruebas estandarizadas (por ej.: evaluar escritura, creatividad, competencias socio-afectivas); informes de autoevaluación de los equipos escolares; sistemas de inspección escolar con reportes cualitativos; uso de indicadores de procesos y encuestas de satisfacción a estudiantes y apoderados, entre otros (Maroy & Voisin, 2015; Ozga, 2014).

En el caso chileno, desde los años 80, se fue gestando un mercado escolar performativo, que refiere a un modelo donde las escuelas -públicas y privadas- deben competir y promocionarse para atraer mayor matrícula (ligado a un financiamiento por estudiante) y, a su vez, el Estado juega un rol de “control a la distancia”, por medio de un conjunto de dispositivos de gobierno, por ejemplo,

fijación de metas e indicadores de logro, exámenes nacionales, clasificación de la calidad de las escuelas, incentivos y sanciones, entre otros (Falabella, 2021). Este enfoque se comenzó a configurar con la creación del Sistema de Medición de la Calidad de la Educación (SIMCE) en 1988, en el contexto de una reforma global de mercantilización de la educación. Luego, desde principios de los 2000 entraron con fuerza las políticas de rendición de cuentas con altas consecuencias, modelo que se consagró mediante la Ley No. 20.529 de 2011 que crea el Sistema Nacional de Aseguramiento de la Calidad de la Educación (SNAC).

En este contexto, y bajo la preocupación de lograr una “evaluación integral”, no reducida a las asignaturas evaluadas en el SIMCE, la ley General de la Educación (Ley No. 20.370 de 2009) estipula que la evaluación de la calidad de la educación escolar incluyese, además de las pruebas SIMCE, el uso de los “Otros Indicadores de la Calidad”, nombrados de este modo en la ley y actualmente denominados desde los organismos públicos como “Indicadores de Desarrollo Personal y Social” (IDPS). Estos son definidos como “un conjunto de índices que entregan información sobre aspectos relacionados con el desarrollo social y personal de los estudiantes de manera complementaria a la información proporcionada por los resultados SIMCE y a los Estándares de Aprendizaje” (Agencia de la Calidad de la Educación, 2015a, p. 10).

La medición de los IDPS es especialmente relevante, pues es contemplada en la metodología estandarizada de “ordenamiento” de la calidad de los establecimientos educacionales, cuya categorización define consecuencias como es, incluso, la revocación del reconocimiento oficial para aquellos establecimientos que obtengan desempeño “insuficiente” durante cinco años. Por lo demás, en un contexto en que la prueba SIMCE ha recibido múltiples críticas (Pino et al., 2016), una encuesta reciente a los directivos de establecimientos escolares muestra que los IDPS es la evaluación gubernamental más valorada (Ramos, 2021). No obstante, existe escaso conocimiento respecto a la evaluación de los IDPS y hay exigua investigación al respecto.

A partir de este escenario, el propósito de este estudio es indagar, por una parte, cómo se gestó la búsqueda por una visión “amplia” de la calidad que derivó en la creación de los IDPS, y, por otra, estudiar de qué modo estos instrumentos -y sus resultados- aportan a una evaluación integral de la calidad. El artículo se divide en tres secciones, la primera expone la metodología utilizada en el estudio; en la segunda se presentan los resultados del estudio distribuidos en: i) la descripción de antecedentes

históricos nacionales e internacionales que representan precedentes e influencias en la gestación de los IDPS; ii) las características técnicas, metodológicas y legales, detallando las dimensiones evaluadas, criterios usados y mecanismos de medición; y iii) la presentación de los principales puntos críticos que surgen desde su implementación. Por último, este artículo devela las inconsistencias técnico políticas de estos indicadores y problematiza sus limitaciones para conseguir una “evaluación integral” dentro de un modelo de rendición de cuentas con altas consecuencias.

Método

Con el propósito de examinar en profundidad la evaluación IDPS y el contexto en que surgió, se realizó un estudio basado en el análisis documental (Priori, 2003). Para ello, en primer lugar, se seleccionaron textos públicos del Estado chileno que hacen referencia a los IDPS (u “Otros indicadores de la calidad”). En segundo lugar, con el fin de comprender los antecedentes históricos de esta evaluación, se recopilaron documentos oficiales relacionados a la evaluación nacional del sistema escolar desde 1981, período de la dictadura militar en que se diseñó la Prueba de Evaluación del Rendimiento Escolar (PER) que antecedió al SIMCE, hasta 2020, año en que cerró el proyecto de investigación. En tercer lugar, con el objetivo de indagar en los discursos de los organismos internacionales, se revisaron informes de las principales agencias que eran citadas en los documentos gubernamentales, estos son: Banco Mundial, BID, OCDE, PREAL, PNUD, UNESCO, UNICEF. En cuarto lugar, se recopilaron los estudios nacionales que hacen referencia a los IDPS. Por último, se revisó la prensa escrita de El Mercurio y La Tercera desde 1973 hasta 2020.

La búsqueda de documentos se llevó a cabo en la Biblioteca Nacional, la biblioteca del Congreso Nacional, la biblioteca del Centro de Perfeccionamiento, Experimentación e Investigaciones Pedagógicas (CPEIP), y la biblioteca del Ministerio de Educación (MINEDUC); y en las páginas web de la Agencia de la Calidad de la Educación (en adelante, Agencia), el MINEDUC, el Consejo Nacional de Educación (CNED), y de los organismos internacionales seleccionados. Las palabras claves de búsqueda fueron: evaluación escolar, calidad integral, evaluación socio-afectiva, “Indicadores de desarrollo personal y social”, “Otros indicadores de la calidad”, Chile. Se excluyó la documentación referida a la evaluación en educación inicial y superior.

En total se seleccionaron 163 documentos. En particular: 10 leyes y decretos, 80 documentos gubernamentales (presentaciones de resultados, informes, estudios, actas

de reunión), 55 documentos internacionales (informes, estudios, revisiones de política), 18 investigaciones nacionales, y diversas notas de prensa (ver Tabla 1).

El análisis de la información cualitativa se realizó mediante la técnica de análisis de contenido (Cisterna, 2005; Piñuel, 2002). Este análisis involucró un examen inductivo de lectura, sistematización y codificación abierta del material, y el análisis de las relaciones establecidas entre categorías. Además, se elaboró una línea del tiempo que permitió configurar la genealogía de los IDPS y las evaluaciones asociadas a ella. Cabe señalar que este estudio se focaliza en analizar el contexto, diseño y aporte de los IDPS a una evaluación integral, pero omite la visión de los actores educativos a nivel local respecto a su uso, lo que podría ser considerado en futuras investigaciones.

Resultados

Antecedentes históricos: La búsqueda por una evaluación integral de la calidad

Las primeras pruebas estandarizadas para evaluar aprendizajes de los estudiantes del sistema escolar chileno fueron la Prueba Nacional de Habilidades Verbales y Matemáticas (1967), y luego el Programa de Evaluación del Rendimiento Escolar (PER) en 1982. Ambas pruebas buscaron contribuir al sistema escolar con información evaluativa “científica” para orientar las políticas públicas y las decisiones pedagógicas (Falabella & Ramos, 2019).

En el PER se vislumbra una primera preocupación por evitar reducir la evaluación nacional al ámbito “cognitivo”, al incorporar indicadores referidos a la autonomía, autoestima, autoevaluación académica, actitud hacia la escuela y el estudio, orientación social, y valores sociales y científicos; y para el caso de octavo básico, orientación vocacional y hacia la continuidad de estudios (Himmel & Majluf, 1983; Himmel, 1997).

La reforma estructural de los 80, bajo la dictadura cívico-militar de Augusto Pinochet, reorganizó el sistema educacional basándose en la subsidiariedad del Estado e introducción de mecanismos de financiamiento mercantil. Con ello, contar con una prueba que permitiera, además, objetivar y comparar la calidad de los colegios resultaba fundamental para orientar la elección de apoderados, dinamizando así la competencia entre escuelas y optimizando la eficiencia del sistema en su conjunto (Prieto, 1983). Tras la suspensión del PER, dicho anhelo se materializó con la creación del SIMCE en 1988, y se hizo obligatoria su aplicación y publicación en la Ley Orgánica Constitucional de la Educación (LOCE, 1990).

El interés embrionario expresado en el PER por una evaluación que trascendiera a lo meramente cognitivo fue reflejado en los primeros años de aplicación del SIMCE, al incorporar una serie de indicadores denominados “estimadores de la calidad educativa” en sus informes, los que incluyeron las variables socioafectivas antes mencionadas y otras de “eficiencia escolar” (tales como retención y egreso oportuno). Durante la década del 90, otras dimensiones fueron contempladas con oscilación, entre ellas, “desarrollo de la creatividad” (1995-1996), “estrategias de aprendizaje” (1993-1995) y “actitudes hacia el ambiente” (1993-1996). Sin embargo, posteriormente “al no tener continuidad ni un marco de referencia común entre años [estas] fueron perdiéndose en el tiempo [así, a fines de esta década] la medición se concentra en los resultados de aprendizaje, abandonando la medición de variables socio-afectivas” (Cariola et al., 2008, pp. 67-68).

En esta década surgió la primera política de rendición de cuentas asociada a incentivos estatales a través de la creación del Sistema Nacional de Evaluación del Desempeño Docente (SNED). Este, tras el PER y el SIMCE, es identificable como el tercer gran instrumento de apoyo a la política educacional que incorporó criterios diversos para definir la “excelencia académica”, considerando: avances en resultados educativos; retención de estudiantes; capacidad del establecimiento para incorporar innovaciones; participación de la comunidad educativa en el desarrollo del proyecto del establecimiento, entre otros.

A fines de los 90 y principios de los 2000, la ingente presión mediática producto del discurso del “estancamiento de resultados SIMCE” y deficientes puntajes en pruebas internacionales, como TIMSS y PISA (Agencia, 2018), generó cuestionamientos sobre la eficacia de las políticas escolares. Por otra parte, emergieron críticas incipientes a una concepción de la calidad centrada en los resultados académicos y a la publicación de rankings SIMCE (Falabella & Ramos, 2019). Tanto la trayectoria previa como estos nuevos estímulos críticos incentivaron la búsqueda de una evaluación que integre más componentes que los meramente cognitivos, lo cual se aprecia en distintos documentos e informes. Así, por ejemplo, la “Comisión SIMCE” (2003), encargada por el MINEDUC de evaluar dicha prueba, reconoció efectos indeseados y limitaciones del SIMCE. En ese marco, vislumbraron como deseable:

obtener una visión completa del sistema educativo incluyendo no sólo resultados de aprendizaje, sino otro tipo de variables, tales como insumos, procesos, impacto laboral o económico, etc. [lo que requiere] obtener y comunicar al país una visión más completa y compleja de la educación (p. 47)

Así, desde 2006, pese al reconocimiento de efectos indeseados, aumentaron los niveles y áreas evaluadas mediante pruebas SIMCE, arguyendo que este incremento permitiría enriquecer la concepción de calidad, evitando el reduccionismo curricular, entregar mejor información a las familias, y hacer un seguimiento a diversas cohortes.

En el plano internacional, aunque distintos organismos han asumido grados disímiles de adherencia y promoción respecto al uso de pruebas estandarizadas, estos han tendido a concordar que es necesario “ampliar el concepto de calidad”. Por ejemplo, durante los 2000 un informe de la OCDE (2004) planteaba:

¿Son los resultados de aprendizaje indicadores completos y suficientes si se quiere tener una visión integral de la calidad educativa? ¿No se está acaso tomando una parte por el todo? ¿Son los resultados de aprendizaje una fotografía exhaustiva de la calidad o se necesita considerar otros indicadores? Si es así, entonces una medición de la calidad requiere ser complementada por evidencia sobre calidad no sólo sobre resultados de aprendizaje, sino también sobre coberturas, equidad, condiciones contextuales para aprender, prácticas docentes y oportunidades de aprendizaje. (p. 95)

Lo planteado era consecuente con una agenda global emergente que otorgaba preponderancia a la calidad de la enseñanza y una ampliación respecto al cómo concebirla y evaluarla. Asimismo, UNICEF (2000) instó la creación de programas que incorporasen una visión integral de la calidad, considerando procesos, entorno y resultados, contemplando no sólo la adquisición de conocimiento, sino también el desarrollo de habilidades, actitudes y una participación ciudadana en la sociedad. La UNESCO (2005) también matizaba el concepto de calidad educativa, y recomendaba utilizar “otras medidas de aproximación del aprovechamiento escolar de los educandos (...) que pueden incluir los beneficios generales para la sociedad” (p. 41). Igualmente, en 2005, por iniciativa de las Cumbres de las Américas, se realizó el II Foro Hemisférico de la Calidad de la Educación, donde ministros de Educación de la región acuerdan “la necesidad de adoptar una visión ampliada de la calidad de la educación (...) cuyo alcance no puede reducirse exclusivamente al ámbito de los logros académicos” (UNESCO, 2009, p. 11).

Paralelamente, en Chile –producto de las críticas del ‘Movimiento Pingüino’ (liderado por estudiantes secundarios)- la Presidenta Bachelet crea el Consejo Asesor Presidencial para la Calidad de la Educación (2006), cuyo informe final reconoció deficiencias en los sistemas de responsabilización, tales como el producir estrechamiento curricular, aspecto que consideraron

podía ser paliado -en coherencia con las recomendaciones del Banco Mundial- a través de la inclusión de otros criterios evaluativos que incluyesen la “calidad de los procesos”, indicadores de cobertura y gestión de los recursos (World Bank, 2007, p. 92).

La clasificación de escuelas elaborada por la Subvención Escolar Presencial (Ley SEP No. 20.248 de 2008) resulta otro antecedente relevante para los IDPS, dado que esta consideró, además de resultados SIMCE, otros indicadores complementarios, tales como: tasas de aprobación y retención; mejoramiento; participación de profesores, padres y apoderados; y la evaluación docente en establecimientos municipales.

Posterior a la ley SEP, se aprueba la Ley General de Educación (LGE) en 2009, que consolidó un Estado evaluador y un sistema de rendición de cuentas con altas consecuencias, con la promesa de garantizar la calidad educativa. Con ello se crea una nueva infraestructura institucional a cargo de evaluar, inspeccionar, sancionar y fiscalizar los establecimientos educativos, por medio de la Agencia de la Calidad y la Superintendencia de la Educación.

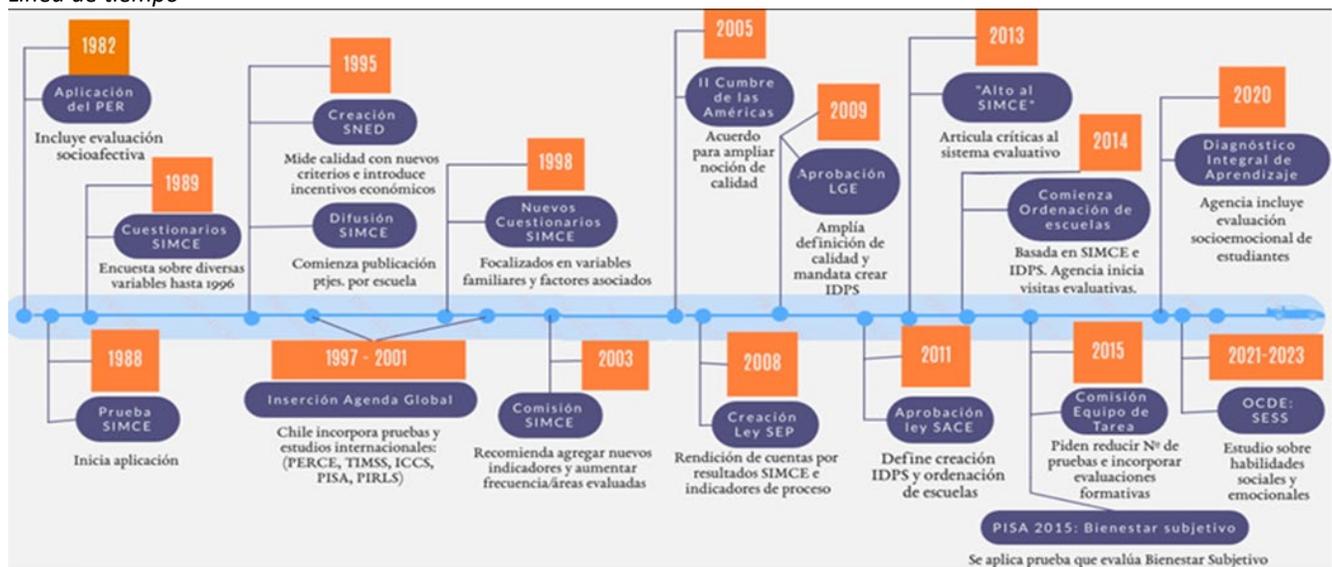
En este marco, la LGE encomienda una “medición integral” tras un extenso debate sobre el uso de las evaluaciones, modo de incorporarlas en el ordenamiento de los establecimientos y las consecuencias asociadas (Congreso Nacional, 2009). Finalmente, se acordó la inclusión de “Otros indicadores de calidad de procesos relevantes de los establecimientos que permitan realizar

una evaluación integral según los objetivos generales establecidos en la ley” (Ley No. 20.370, 2009, art. 38); a lo que se sumó las “Visitas de Evaluación y Orientación del Desempeño” de la Agencia a establecimientos educativos, basadas en los Estándares Indicativos de Desempeño.

Posteriormente, en 2011 se crea la Ley del Sistema Nacional de Aseguramiento de la Calidad en Educación (SNAC), la que generó “un marco normativo que define la calidad de la educación de manera integral, entendiendo la calidad educativa en un sentido amplio (...) yendo más allá de los aprendizajes académicos” (Agencia, 2015b, p. 9). En la práctica, esto se establece mediante una nueva ordenación que crea cuatro categorías de desempeño para todos los establecimientos del país (SNAC, 2011, art. 112), clasificándolos en base a los resultados de aprendizajes obtenidos en las pruebas SIMCE (cuya ponderación mínima por ley es del 67% en la metodología de ordenación), medidas de progreso (tendencia SIMCE) y los ocho IDPS que evalúan aspectos sociales y personales del estudiante (Ley No. 20.529 de 2011, art.12 y art.17), corrigiendo luego los puntajes en base a las características socioeconómicas de los estudiantes por medio de una fórmula definida por la Agencia (Agencia, 2014; MINEDUC, 2014).

Desde la aprobación de la Ley SNAC existe un giro discursivo respecto a la noción de la calidad. Mientras en los años 90 y principios de los 2000 se relevó la idea de “más y mejores aprendizajes”, traducido en resultados

Figura 1
Línea de tiempo



estandarizados, en una fase más reciente surge la narrativa del “equilibrio entre presión y apoyo” y la búsqueda por una evaluación integral “más allá del SIMCE” (ver Figura 1).

Desde la Agencia, Carlos Henríquez siendo Secretario Ejecutivo de ella, incorporó este ideario, liderando una estrategia que apaciguó las críticas al sistema evaluativo (ACER, 2013; Flórez, 2013; Ortiz, 2012) atribuyendo a la incorporación de los IDPS un cambio de paradigma evaluativo.

Calidad es mucho más que el SIMCE (...). Antes era sólo SIMCE, hoy día son los Indicadores de Desarrollo Personal y Social, con una señal clara de que importa un desarrollo integral en los colegios. Antes era sólo puntaje, hoy día es puntaje, estándares y eje de contenidos (...) Dimos una señal muy clara con este nuevo modelo de evaluación (...) más que buscar resultados, mirar procesos en el aula, orientando, retroalimentando (...) Estamos promoviendo una cultura distinta. (Agencia Educación, 2016)

Los cuestionamientos tuvieron cabida en el programa presidencial del segundo gobierno de Bachelet (2013), donde afirmaba que se debía “impulsar una comprensión más compleja e integral de calidad para recuperar el sentido de la labor educativa” (p. 16).

En 2015 se conformó una comisión revisora del sistema evaluativo llamada “Equipo de Tarea” (2015) que aspiraba a crear un “sistema equilibrado e integral”. Esta comisión sugirió la reducción en la cantidad de pruebas censales aplicadas anualmente, evitar rankings SIMCE, y agregar pruebas estandarizadas voluntarias en los niveles en que no se aplica el SIMCE para que entreguen información a los equipos escolares. Posteriormente esta última recomendación se tradujo desde el año 2016 en la llamada “evaluación progresiva”. No obstante, a pesar de posibles mejoras, se mantuvo el enfoque global del modelo evaluativo y se omitió un cuestionamiento estructural respecto a la utilidad, uso y consecuencias asociadas al sistema vigente (Flórez & Oyarzún, 2016).

También es importante notar que en 2017 la Agencia creó una campaña abierta en Twitter denominada “#CalidadEs”, difundiendo imágenes con diversos apelativos junto a este concepto, tales como “confianza”, “diversidad” y “equidad de género”, consultando opiniones respecto a qué era calidad para cada persona. Era la relativización de la noción de calidad que se contradice con la visión positivista de calidad entendida como una cualidad medible, comparable, objetiva y transparente. Adicionalmente, en el contexto de la pandemia COVID, se creó el Diagnóstico Integral de

Aprendizaje (DIA), instrumento voluntario cuya finalidad es evaluar el desempeño en lectura, matemáticas y el estado socioemocional de los estudiantes.

En cuanto al debate a nivel global, si bien desde mediados de 1960 iniciativas como la *Six Subject Survey* de la International Association for the Evaluation of Educational Achievement (IEA) mostraban interés por aspectos no cognitivos de la formación (Falabella & Ramos, 2019), esta tendencia ha cobrado particular importancia en las últimas décadas, expresada en múltiples publicaciones (OCDE, 2017; PNUD, 2012; PNUD & UNICEF, 2014) que han planteado la preocupación por cautelar el “bienestar subjetivo” de los alumnos en los establecimientos escolares. Esto se vincula con, por ejemplo, la capacidad de los estudiantes de crear un proyecto de vida propio, generar vínculos significativos, desarrollar sentido de pertenencia y bienestar en el espacio escolar, participar e influir en la sociedad, entre otros.

Ello derivó en la creación de la prueba PISA de Bienestar Subjetivo, aplicada en 2015, hito relevante en cuanto busca medir aspectos que trascienden a lo cognitivo (OCDE, 2018), y recientemente el Estudio Internacional de Competencias Socioemocionales (SSES), que busca recabar evidencia empírica sobre las habilidades emocionales y sociales de estudiantes (10-15 años) y su importancia para la vida. En definitiva, el anhelo por una evaluación amplia e integral a nivel nacional, más que una preocupación contingente de los IDPS, responde a una inquietud consistente durante las últimas décadas.

Marco normativo y técnico para la creación de los IDPS

La ley SNAC (2011) detalla las exigencias para la elaboración de los IDPS, mandatando que deben: i) ser evaluados mediante instrumentos estandarizados, válidos, confiables, objetivos y transparentes; ii) aplicarse obligatoriamente a todos los colegios con reconocimiento oficial del Estado; iii) ser renovados cada seis años; iv) proveer los resultados de la medición como información relevante, de fácil comprensión para la comunidad y comparable en el tiempo; v) ser considerados en la ordenación de los establecimientos; vi) y poseer una ponderación inferior al 33% del total en dicha ordenación, al considerarse los resultados de los Estándares de Aprendizaje en el otro 67%; y vii) procurar ser coherentes con los principios desarrollados en la LGE (derecho a la educación, libertad a la enseñanza, calidad, equidad, transparencia, entre otros). Esta ley, además, obliga a cautelar la validez, confiabilidad, objetividad, y estandarización de los instrumentos utilizados. Para esto, los IDPS deben ser censales (art. 17), comparables en el tiempo (art. 20) y transparentes.

El proceso de selección y definición de los IDPS se inició en 2011 en una comisión del MINEDUC, y fueron aprobados por el CNED en 2013, incorporando como indicadores sólo aquellos que lograron consenso, descartando “resultados PSU”, “éxito oportuno”, “repetencia” y “satisfacción de padres y apoderados”, por los efectos indeseados que podrían conllevar (CNED, 2013a; MINEDUC, 2013a).

En cuanto al proceso de validación de la propuesta, ésta fue compartida con “expertos en educación”, sumado a sostenedores y directores de establecimientos, quienes manifestaron su preocupación “por el reporte de los Otros Indicadores de Calidad como resultados brutos, aparte de los resultados de la Ordenación general”, al poseer los IDPS una baja ponderación en relación a la medición de logros de aprendizaje, transmitiendo con ello que “lo central seguían siendo los resultados académicos” (MINEDUC, 2013b, p. 74). También cabe agregar que durante el proceso de consultas surgieron reparos dentro del propio Consejo Ejecutivo de la Agencia ante “su preocupación por los indicadores que se miden a través de los cuestionarios autoreportados [de parte de apoderados y alumnos], dentro de un sistema de evaluación que tiene consecuencias para los establecimientos”, lo que fue refutado por el comité interno a cargo, quienes aseguraron que “los indicadores medidos a través de cuestionarios eran medidos mundialmente con instrumentos de este tipo [...] y que se habían realizado los análisis (...) para asegurar la validez y confiabilidad del instrumento” (MINEDUC, 2013b, p. 74).

Finalmente, los ocho indicadores zanjados fueron: i) autoestima académica y motivación escolar; ii) clima de convivencia escolar; iii) participación y formación ciudadana; iv) hábitos de vida saludable; v) asistencia escolar; vi) retención escolar; vii) equidad de género; y viii) titulación técnico profesional. Para la medición de los cuatro primeros, se utilizan “Cuestionarios de Calidad y

Contexto de la Educación” (aplicados durante el SIMCE). Estos son contestados por estudiantes, padres, apoderados y/o docentes, según el indicador. Asimismo, el MINEDUC optó por la modalidad de cuestionarios autoreportados para evaluar los indicadores de percepción y actitudinales argumentando que “el modo en que los individuos perciben su ambiente es un predictor sólido del desempeño social, emocional y académico, incluso de mayor peso que lo que realmente sucede” (MINEDUC, 2013b, p. 51). Por su parte, esgrimieron que las actitudes -entendidas como la predisposición a actuar de manera consistente con respecto a un determinado objeto o situación- contemplan una visión más integral del aprendizaje al incluir componentes afectivos, cognitivos y conductuales.

Respecto a los otros cuatro indicadores, estos son calculados a través de un análisis factorial exploratorio de cada dimensión, grado y actor evaluado (en base al subconjunto de ítems que componen los IDPS), y el uso de un análisis factorial con rotación (permite asociación entre factores principales), cuya fiabilidad se mide mediante el coeficiente Alfa de Cronbach y el índice KMO. Así, por ejemplo, los puntajes de la dimensión “hábitos alimentarios” no fueron publicados, “debido a la baja consistencia interna que se obtuvo en 2014 (Alfa de Cronbach <0.7)” (Agencia, s/fa, p. 24).

En dicho modelo multidimensional consideraron variables dependientes (resultados SIMCE e IDPS) e independientes asociadas al desempeño de los estudiantes, entre ellas: características de los estudiantes y sus familias (género del estudiante, características del origen social tales como el nivel educacional alcanzado por los padres, el ingreso familiar, expectativas y prácticas de su entorno familiar, y el involucramiento parental), y de sus establecimientos (características sociodemográficas como la dependencia administrativa, la modalidad educativa, el género del curso y el tamaño

Tabla 1

Caracterización de los instrumentos de evaluación de los IDPS autoreportados

| Indicador | Dimensión | Actor | Medición |
|--|---|---|---------------|
| 1. Autoestima académica y motivación escolar | Autopercepción y autovaloración académica | Estudiantes | Cuestionarios |
| | Motivación escolar | | |
| 2. Clima de convivencia escolar | Ambiente de respeto | Estudiantes, Profesores, Aporerados | |
| | Ambiente organizado | | |
| | Ambiente seguro | | |
| 3. Participación y formación ciudadana | Sentido de pertenencia | Estudiantes, Aporerados | |
| | Participación | | |
| | Vida democrática | | |
| 4. Hábitos de vida saludable | Hábitos alimentarios | Estudiantes | |
| | Hábitos de vida activa | | |
| | Hábitos de autocuidado | | |

Fuente: Elaboración propia en base a documentos MINEDUC (2013b, 2014b).

del establecimiento o número de matriculados). Las Tablas 1 y 2 detallan las dimensiones evaluadas por indicador, los actores consultados en su medición y cómo se evalúan.

Los índices de los IDPS poseen una escala de 0 a 100, donde a mayor índice mejor se encuentra el indicador. Sin embargo, cabe aclarar que estos son incomparables entre indicadores y entre grados; los resultados solo pueden compararse entre años distintos para un mismo indicador de un mismo grado. Por ello, la Agencia exige mayor cantidad de mediciones para ver tendencias más claras (Agencia, 2015b).

Análisis crítico de los IDPS y sus resultados

Desde 2014, la Agencia ha difundido anualmente los puntajes IDPS, buscando ampliar una comprensión más compleja sobre la calidad. Sin embargo, el uso mediático de los IDPS y sus resultados no se encuentran exentos de cuestionamientos; a continuación, se plantean siete puntos críticos. Un primer reparo se da por el mecanismo utilizado para recolectar la información evaluativa en cuatro de los IDPS: los cuestionarios autoreportados. Estos son presentados como una evaluación externa, estandarizada y objetiva, cuando se basan en efecto en una encuesta de percepción de estudiantes, apoderados y profesores respecto a las dimensiones evaluadas.

El segundo es que los resultados de las dimensiones evaluadas por medio de los cuestionarios tienden a

arrojar un alto nivel de aprobación, con una baja dispersión en la distribución de las respuestas, lo que merma la validez de los resultados. Por ejemplo, a nivel nacional, en el año 2016 los resultados en tres de los cuatro indicadores fluctuaron entre un 97% y un 99% en el nivel de logro alto y medio; mientras “hábitos de vida saludable”, indicador que tuvo el puntaje más bajo, obtuvo un 88% (Agencia, s/fb). La misma Agencia reconoce en diversos informes que, pese a leves diferencias en algunas dimensiones y grados, mayoritariamente los indicadores evaluados no presentan “diferencias entre los grupos respecto a la media nacional en ninguna de las mediciones” (2015a, p. 60), ni tampoco tendencias claras que establezcan diferencias significativas según localización regional, “ni por GSE y dependencia administrativa de los establecimientos ni por sexo de los estudiantes” (Agencia, 2015c, p. 75; ver como ejemplo la Figura 2).

Por ende, si bien la Agencia afirma poseer niveles de validez y confiabilidad en base a los análisis factoriales que realiza, la falta de dispersión mostrada devela que, en la práctica, puede haber importantes sesgos. Aquello, especialmente referido a una alta deseabilidad en las respuestas, al tratarse de una evaluación con consecuencias para los establecimientos. De hecho, la Agencia reconoce la complejidad que representa “evitar insinuar o revelar la expectativa o la respuesta deseada de las preguntas e ítems” (s/fb, p. 56), visibilizando el riesgo de deseabilidad que produce en los encuestados.

Tabla 2

Caracterización de los IDPS elaborados a partir de Registros del MINEDUC y la Agencia

| Indicador | Definición | Fuente | Categorías |
|-----------------------------------|---|--|---|
| 5. Asistencia escolar | Número de asistencia escolar diario | Acta de Registro de Calificaciones Finales y Promoción Escolar en el Sistema General de Información de Estudiantes (SIGE) | Si más del 60% de los alumnos presenta una asistencia destacada o normal, se presenta una asistencia escolar ‘adecuada’; si el 40% o más presenta inasistencia reiterada o grave, se considera como ‘con problemas’ |
| 6. Retención Escolar | Porcentaje de estudiantes por establecimiento que permanece en el sistema escolar durante un año | Información anual reportada a nivel individual por estudiante por el SIGE | En caso de tener un 2% o más de estudiantes que desertan del sistema se estima ‘con problemas’ |
| 7. Equidad de Género | Logro equitativo de resultados entre hombres y mujeres en colegios mixtos Resultados SIMCE en Lectura y Matemática | A partir del cálculo de la brecha entre hombres y mujeres en las pruebas SIMCE, se asigna una categoría Baja, Regular, Adecuada, expresada en puntajes | |
| 8. Titulación Técnico Profesional | Proporción de estudiantes que reciben título de técnico nivel medio tras egresar de la EMT | Datos entregada por Secretarías Regionales Ministeriales de Educación | En caso de ser mayor al 70% que se titula de la especialidad TP, se le considera ‘adecuada’; de lo contrario, ‘con problemas’. |

Fuente: Elaboración propia en base a documentos MINEDUC (2013b, 2014b).

Este problema fue además previsto tanto por el CNED como por el comité de expertos consultado previo a su aprobación (CNED, 2013b).

Tercero, son exiguas las experiencias en el mundo que evalúan este tipo de variables - como las definidas en los IDPS - utilizando cuestionarios (a excepción del indicador “participación y formación ciudadana”). La Agencia (2013b) reconoce que “en general, en los sistemas de responsabilización con consecuencias no se evalúa de manera censal el desarrollo de las “habilidades no cognitivas” (p. 50), privilegiándose el uso de visitas de evaluación a los establecimientos y elaboración de informes cualitativos para evaluar la dimensión personal y social. Con esto, se visibiliza la falta de experiencias comparadas que puedan aportar evidencia en torno a los beneficios de esta política atípica en el mundo.

Cuarto, se cuestiona si la intervención y quehacer de la escuela puede efectivamente influir en la mejora y desarrollo de aquellas dimensiones personales y sociales por las que son evaluadas y categorizadas. En ese sentido, estudios etnográficos encargados por la misma Agencia han visibilizado tensiones relevantes. Uno de ellos puntualizó “la ausencia de prácticas institucionalizadas para la promoción del desarrollo socioemocional” (Agencia, 2015c, p. 19) en las escuelas, mientras que otra investigación cualitativa focalizada en ‘hábitos de vida

saludable’ visibilizó que, si bien los colegios pueden emprender acciones para fomentarlos, estos reconocen dificultades que trascienden a su margen de acción, como las limitaciones económico-contextuales, expresadas en la falta de “acceso a una oferta saludable” (p. 36), y el rol de las familias, el cual “marca diferencias importantes entre escuelas [...] y cuya ausencia] para fomentar un consumo balanceado (...) dificulta su promoción” (Torres et al., 2019, p. 385).

Quinto, se abre un debate en torno a la selección de estos indicadores en la medida que excluyen aspectos relevantes, como son, por ejemplo, el discernimiento ético y conciencia medioambiental, por nombrar algunos. Asimismo, cabe cuestionar los reduccionismos conceptuales en los que incurren los actuales IDPS, como acontece con el indicador ‘equidad de género’, donde se asimila su evaluación a la mera diferenciación de puntajes SIMCE entre hombres y mujeres, obviando el impacto - más allá de ‘logros de aprendizaje’- de los micromachismos y estereotipos de género presentes tanto en el currículum como en las prácticas docentes (Azúa et al., 2019; Palestro, 2016).

Lo mismo ocurre en otros indicadores, como, por ejemplo, en el de ‘hábitos de vida saludable’, cuya dimensión de ‘autocuidado’ se reduce a una definición centrada en las actitudes, conductas y percepciones

Figura 2
Resultados IDPS medidos por cuestionarios y presentados por GSE

| Indicador | GSE | 2014 | | | | 2015 | | | | 2016 | | | 2017 | | | 2018 | | |
|---|------------|------|----|----|----|------|----|----|----|------|----|----|------|----|----|------|----|----|
| | | 4b | 6b | 8b | II | 4b | 6b | 8b | II | 4b | 6b | II | 4b | 8b | II | 4b | 6b | II |
| Autoestima académica y motivación escolar | Bajo | 74 | 75 | 74 | 73 | 75 | 75 | 75 | 73 | 74 | 74 | 73 | 74 | 76 | 74 | 74 | 75 | 74 |
| | Medio bajo | 74 | 74 | 73 | 73 | 74 | 74 | 74 | 74 | 74 | 74 | 74 | 75 | 75 | 74 | 75 | 75 | 74 |
| | Medio | 74 | 74 | 74 | 74 | 74 | 74 | 74 | 74 | 74 | 74 | 74 | 74 | 74 | 74 | 74 | 74 | 73 |
| | Medio alto | 74 | 74 | 74 | 75 | 74 | 74 | 73 | 74 | 74 | 74 | 74 | 74 | 72 | 74 | 74 | 73 | 73 |
| | Alto | 73 | 73 | 76 | 76 | 74 | 74 | 74 | 76 | 75 | 74 | 75 | 74 | 73 | 75 | 74 | 73 | 74 |
| Clima de convivencia escolar | Bajo | 72 | 74 | 75 | 72 | 74 | 75 | 75 | 72 | 74 | 75 | 72 | 73 | 76 | 74 | 74 | 76 | 75 |
| | Medio bajo | 72 | 73 | 74 | 74 | 73 | 73 | 74 | 74 | 73 | 73 | 74 | 74 | 75 | 75 | 74 | 75 | 75 |
| | Medio | 75 | 76 | 76 | 77 | 76 | 75 | 76 | 77 | 76 | 76 | 77 | 76 | 75 | 76 | 75 | 75 | 75 |
| | Medio alto | 77 | 78 | 76 | 77 | 77 | 77 | 76 | 77 | 77 | 76 | 76 | 77 | 75 | 76 | 76 | 75 | 74 |
| | Alto | 79 | 79 | 78 | 79 | 78 | 78 | 78 | 79 | 78 | 77 | 78 | 78 | 76 | 77 | 77 | 76 | 76 |
| Participación y formación ciudadana | Bajo | 77 | 79 | 78 | 76 | 77 | 79 | 78 | 76 | 78 | 79 | 76 | 78 | 79 | 77 | 79 | 80 | 78 |
| | Medio bajo | 76 | 77 | 76 | 76 | 77 | 77 | 76 | 77 | 78 | 78 | 77 | 78 | 78 | 77 | 78 | 78 | 77 |
| | Medio | 77 | 78 | 77 | 79 | 78 | 78 | 77 | 79 | 78 | 77 | 78 | 78 | 77 | 77 | 77 | 77 | 76 |
| | Medio alto | 78 | 78 | 78 | 79 | 79 | 78 | 79 | 79 | 78 | 77 | 78 | 78 | 75 | 77 | 77 | 76 | 76 |
| | Alto | 80 | 80 | 81 | 82 | 80 | 80 | 81 | 82 | 79 | 78 | 80 | 78 | 77 | 79 | 78 | 76 | 78 |
| Hábitos de vida saludable | Bajo | 69 | 71 | 74 | 69 | 69 | 71 | 72 | 70 | 68 | 70 | 70 | 69 | 73 | 70 | 69 | 71 | 70 |
| | Medio bajo | 69 | 69 | 70 | 68 | 69 | 69 | 70 | 69 | 69 | 69 | 70 | 69 | 70 | 69 | 69 | 70 | 70 |
| | Medio | 70 | 70 | 69 | 70 | 70 | 69 | 69 | 70 | 71 | 70 | 70 | 70 | 69 | 70 | 70 | 70 | 69 |
| | Medio alto | 71 | 70 | 68 | 70 | 71 | 70 | 69 | 69 | 71 | 70 | 69 | 71 | 69 | 70 | 71 | 70 | 70 |
| | Alto | 71 | 72 | 71 | 75 | 74 | 73 | 71 | 72 | 72 | 72 | 72 | 73 | 73 | 74 | 72 | 71 | 73 |

Fuente: Agencia, 2019 (p. 13)

vinculadas con la sexualidad, el consumo de tabaco, alcohol, drogas, e higiene (MINEDUC, 2013a). Por otra parte, otros indicadores como ‘participación y formación ciudadana’, solamente evalúan las percepciones de los estudiantes en las tres dimensiones -sentido de pertenencia, participación y vida democrática-, mientras apoderados sólo son considerados en ‘participación’, y los docentes -en algo tan relevante como la vida democrática en la escuela- son completamente excluidos, lo que evidencia que en estas definiciones subyacen “determinadas visiones sobre la sociedad [...por lo que] evaluar no se reduce a una práctica técnica, sino que posee un carácter político, en tanto incide directamente en los valores del sistema educativo que queremos” (Pino et al., 2016, p. 348).

Sexto, surge una crítica que refiere a las explicaciones dadas a partir de los resultados. La Agencia (2015b) incluyó en su primera difusión de resultados un análisis en torno a factores asociados a dichos resultados los que proveían “evidencia respecto de aquellos factores que más se relacionan con el desempeño académico, personal y social de los estudiantes, y que la propia escuela efectivamente tiene la posibilidad de modificar” (p. 7). En ese sentido, resaltaron que el indicador de clima de convivencia escolar, y factores como el liderazgo directivo, retroalimentación docente e involucramiento parental a edad temprana “se asocian a importantes diferencias de resultados, pudiendo dar cuenta de hasta 76 puntos [SIMCE] de diferencia, lo que nuevamente resalta la importancia de aquellas prácticas” (Agencia, 2015b, p. 50).

Esta hipótesis erige un nuevo cuestionamiento: constituir una suposición de causalidad no demostrada. Que exista una relación entre convivencia escolar y puntajes SIMCE no implica que la primera sea causa directa de la segunda. Así lo corrobora un análisis realizado por medio del modelo multinivel, que constata la imposibilidad de “establecer relaciones causales, es decir, no permite identificar si el cambio en determinada variable impactaría en los valores que toma otra variable” (Agencia, 2015b, p. 12).

Por último, desde una visión más sistémica, es necesario criticar el uso discursivo atribuido a los IDPS como garantizador de procesos de evaluación integral y problematizar el modelo de rendición de cuentas en el que se insertan. Desde su incorporación en la presentación de resultados anuales, se ha exacerbado comunicacionalmente su contribución para transitar supuestamente hacia una visión amplia de la calidad. Sin embargo, cabe recordar que la ordenación de escuelas le otorga un 67% de ponderación al logro de estándares de aprendizaje, medidos vía SIMCE, lo que sumado a los

indicadores de tendencia le deja sólo un 27% a ser distribuido entre los IDPS. Por lo tanto, en la práctica, dada su baja ponderación y exigua dispersión a través de los años, las escuelas siguen siendo juzgadas y responsabilizadas cuasi exclusivamente por sus resultados SIMCE. Así lo plasma la categorización de desempeño de la Agencia al agrupar el rendimiento de las escuelas según dependencia donde la diferencia es sustantiva. Ello constata que “la categorización de acuerdo a desempeño no está siendo independiente del nivel socioeconómico. El “control” de tal variable no estaría siendo efectivo” (Ramos, 2021, p. 29).

Adicionalmente, existe investigación reciente que problematiza la capacidad efectiva de que las escuelas, en un modelo como el chileno -donde prima la competencia entre establecimientos por atraer matrícula (y, con ello, financiamiento)- puedan lograr una mejora integral y simultánea de sus indicadores, al constatar que la mejora en resultados SIMCE se suele dar descuidando otras dimensiones de la calidad educativa, expresado en la reducción en los puntajes de los IDPS (Gajardo & Grau, 2017).

De esta forma, se erige la ilusión de poseer una evaluación integral con la inclusión de los IDPS mientras persiste la preponderancia que preserva el SIMCE y las múltiples consecuencias pecuniarias asociadas a sus resultados, teniendo un alto impacto en las prioridades de las escuelas y su quehacer cotidiano (Acuña et al., 2014; Falabella, 2020; Flórez Petour et al., 2018).

Discusión y Conclusiones

El intento de evaluar de modo integral la calidad de la educación ha sido un anhelo presente desde el PER en la década de los 80s. Anhelo que se fue sofisticando en el tiempo a partir de diversas leyes (SNED, SEP, LGE, SNAC), imbuidas en discursos e iniciativas a nivel global respecto a las preocupaciones y recomendaciones de “trascender” las pruebas estandarizadas y “ampliar” la evaluación de la calidad educativa.

El caso chileno se mantiene bajo un modelo de rendición de cuentas con altas consecuencias, no obstante, se han incluido nuevos ámbitos e instrumentos, como los IDPS, en el esfuerzo de ampliar el sistema evaluativo, además de otras medidas, tales como las visitas evaluativas de la Agencia, la ampliación de pruebas SIMCE a nuevas áreas de Evaluación, la Evaluación Progresiva y el DIA.

A partir del análisis de la construcción y medición de los IDPS, se hallan inconsistencias metodológicas y de validez con relación a aquellas dimensiones evaluadas. En ese

sentido, el artículo problematiza las limitaciones de los IDPS en el marco de responsabilización con altas consecuencias y la falta de evidencia técnica que respalda la implementación de esta política. Tal como la misma Agencia admite -en un documento de circulación restringida (obtenido por Ley de Transparencia)-, la "ausencia de experiencias [internacionales] en sistemas de rendición de cuentas por resultados educativos en los ámbitos abordados por los indicadores de desarrollo personal y social se constituye como la principal debilidad de esta iniciativa nacional" (Agencia, s/fc, p. 21).

Asimismo, se cuestiona el uso discursivo de los IDPS por parte de la Agencia al comunicar percepciones captadas a través de encuestas como si representasen una medición externa y objetiva de un fenómeno, difundiendo incluso causalidades no demostradas al vincular la mejoría de los IDPS con resultados SIMCE. Junto a ello, le atribuyen un cambio de paradigma en términos evaluativos, sin embargo, en la práctica, las escuelas siguen siendo juzgadas y responsabilizadas primordialmente por logros de aprendizajes medidos vía SIMCE, aspecto que contraviene el propósito planteado por la misma Agencia de una "retroalimentación cualitativa y contextualizada (...) que evalúe procesos y no solo resultados, con mayor contacto presencial" (2015d, p. 22).

En conclusión, en base a lo examinado, se argumenta que la idea de una evaluación equilibrada e integral, representada en los IDPS, es más bien una ilusión, como también una estrategia discursiva, en que sus límites se circunscriben en un insoslayable modelo de evaluación estandarizado y censal con altas consecuencias el que, en su diseño actual, imposibilita el tránsito hacia un nuevo paradigma educacional. Ello requiere indefectiblemente transformar el actual paradigma de evaluación y transitar hacia una evaluación formativa, profesionalizante y participativa.

Referencias¹

Acuña, F., Assaél, J., Contreras, P., & Peralta, B. (2014). La traducción de los discursos de la política educativa en la cotidianeidad de dos escuelas municipales chilenas: La metáfora médica como vía de análisis. *Psicoperspectivas*, 13(1), 46-55. <http://dx.doi.org/10.5027/psicoperspectivas-vol13-issue1-fulltext-363>

Agencia de Calidad de la Educación*. (Agencia, 2013a). *Evaluación de los procesos y productos relacionados con la producción de instrumentos, operaciones de campo y manejo de datos de las*

pruebas nacionales SIMCE. ACER.

Agencia de la Calidad de la Educación*. (Agencia, 2013b). *Fundamentos metodología de ordenación de establecimientos*. <https://bit.ly/3B5mUwP>

Agencia de la Calidad de la Educación*. (Agencia, 2014). *Minuta Informe metodología de clasificación en categorías de desempeño*. <https://bit.ly/3xHuc7F>

Agencia de la Calidad de la Educación*. (Agencia, 2015a). *Reporte de la Calidad: Evolución de los Indicadores de Calidad de la Educación en Chile*. <https://bit.ly/3hCtjrl>

Agencia de la Calidad de la Educación*. (Agencia, 2015b). *Factores asociados a resultados SIMCE e indicadores de desarrollo personal y social 2014*. <https://bit.ly/3xELp1K>

Agencia de la Calidad de la Educación*. (Agencia, 2015c). *Los Indicadores de Desarrollo Personal y Social en los establecimientos educacionales chilenos: Una primera mirada*. <https://bit.ly/3elkiRI>

Agencia de la Calidad de la Educación*. (Agencia, 2015d). *Calidad educativa desde la percepción de los actores clave del sistema*. <https://bit.ly/3hEhZuO>

Agencia de la Calidad de la Educación*. (Agencia, 2016, jun. 15). *Más allá de las notas: ampliando fronteras en evaluación de aprendizajes* [video]. YouTube. <https://bit.ly/3hEljGD>

Agencia de la Calidad de la Educación*. (Agencia, 2018). *Evaluaciones nacionales e internacionales de aprendizaje: Periodo 2004-2016*. <https://bit.ly/3hGaVv8>

Agencia de la Calidad de la Educación*. (Agencia, 2019). *Evaluaciones Nacionales e Internacionales de Aprendizaje. Periodo 2004-2018*. <https://bit.ly/36zF77o>

Agencia de la Calidad de la Educación*. (Agencia, s/fa). *Desarrollo personal y social: Otros indicadores de calidad educativa. descripción, metodología de construcción y resultados de los indicadores construidos en base a los cuestionarios de calidad y contexto de la educación*. <https://bit.ly/2UQq3AO>

Agencia de la Calidad de la Educación*. (Agencia, s/fb). *Informe Técnico 2017: Indicadores de Desarrollo Personal y Social (IDPS) medidos a través de cuestionarios*. <https://bit.ly/2UKiNpg>

Agencia de la Calidad de la Educación*. (Agencia, s/fc). *Evaluación y rendición de cuentas con Indicadores de Desarrollo Personal y Social en el sistema educativo en Chile*. División de Estudios, Agencia de la Calidad.

Au, W. (2007). High-stakes testing and curricular control: A qualitative meta synthesis. *Educational*

¹ El asterisco indica que la fuente fue analizada en este estudio.

- Researcher*, 36(5). 258-267.
<https://doi.org/10.3102/0013189X07306523>
- Azúa, X., Saavedra, P., & Lillo, D. (2019). Injusticia social naturalizada: Evaluación sesgo de género en la escuela a partir de la observación de videos de la evaluación docente. *Perspectiva Educacional*, 58(2), 69-97.
<https://dx.doi.org/10.4151/07189729-vol.58-iss.2-art.916>
- Bachelet, M.* (2013). *Chile de todos: Programa de gobierno Michelle Bachelet 2014-2018*.
<https://bit.ly/3r9CJ0M>
- Cariola, L., Cares, G., & Rivero, R.* (2008). Sistemas de evaluación como herramientas de políticas. *Revista Iberoamericana de Evaluación Educativa* 1(1), 65-75. <https://bit.ly/2VMFKbX>
- Cisterna, F. (2005). Categorización y triangulación como procesos de validación del conocimiento en investigación cualitativa. *Theoria*, 14(1), 61-71.
<https://bit.ly/3hEtZwD>
- Comisión Nacional para la Modernización de la Educación*. (1994). *Los desafíos de la educación chilena frente al siglo 21*. <https://bit.ly/3B73WpE>
- Comisión para el Desarrollo y Uso del Sistema de Medición de la Calidad de la Educación* (2003). *Evaluación de aprendizajes para una educación de calidad*. <https://bit.ly/3hEizUF>
- Consejo Asesor Presidencial*. (2006). *Informe final del consejo asesor presidencial para la calidad de la educación*. <https://bit.ly/2UM4xfw>
- Consejo Nacional de Educación*. (2013b). *Acuerdo No. 22/2013*. <https://bit.ly/36ylZqH>
- Consejo Nacional de Educación*. (CNED, 2013a). *Acuerdo No. 31/2013*. <https://bit.ly/36BY2i8>
- Equipo de Tarea para la Revisión del SIMCE*. (2015). *Hacia un sistema completo y equilibrado de evaluación de los aprendizajes en Chile: Informe Equipo de Tarea para la revisión del SIMCE*. <https://bit.ly/36CTSGA>
- Falabella, A. (2014). The performing school: Effects of market & accountability policies. *Education Policy Analysis Archives*, 22(51), 1-29.
<http://dx.doi.org/10.14507/epaa.v22n70.2014>
- Falabella, A. (2020). The ethics of competition: Accountability policy enactment in Chilean schools' everyday life. *Journal de Education Policy*, 35(1), 23-45.
<https://doi.org/10.1080/02680939.2019.163527>
- Falabella, A. (2021). The seduction of hyper-surveillance: standards, testing, and accountability. *Educational Administration Quarterly*, 57(1), 113-142.
<https://doi.org/10.1177/0013161X20912299>
- Falabella, A., & 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. <https://bit.ly/36HXwPr>
- Flórez Petour, M. T., Rozas Assaél, T., Gysling, J., & Olave Astorga, J. M. (2018). The consequences of metrics for social justice: Tensions, pending issues, and questions. *Oxford Review of Education*, 44(5), 651-667.
<https://doi.org/10.1080/03054985.2018.1500356>
- Flórez, M. T. (2013). *Análisis crítico de la validez del sistema de medición de la calidad de la educación (SIMCE): Informe Final*. Oxford University Centre for Educational Assessment. <https://bit.ly/3i74LpK>
- Flórez, M. T., & Oyarzún, G. (2016). Resultados SIMCE y Plan de Evaluación 2020: Nudos críticos y perspectivas de cambio. *Cuadernos de Educación*, 73, 1-11. <https://bit.ly/3dEckh6>
- Gajardo, F., & Grau, N. (2017). *Competition among schools and educational : Tension between various objectives of educational policy. Serie de Documentos de Trabajo, (445)*. Universidad de Chile. <https://bit.ly/3vS1lxj>
- Himmel, E.* (1997). Impacto de los sistemas de evaluación del rendimiento escolar: El caso de Chile. In B. Álvarez, & M. Ruiz-Casares (Eds.), *Evaluación y reforma educativa: Opciones de política* (125-147). PREAL.
- Himmel, E., & Majluf, N.* (1983). Programa de evaluación de rendimiento escolar en 1982. *Revista de Educación*, 113(22).
- Hursh, D., Deutermann, J., Rudley, L., Chen, Z., & McGinnis, S. (2020). *Opting out: The story of the parents' grassroots movement to achieve whole-child public schools*. Stylus.
- Ley No. 20.248.* (2008, feb. 1). *Establece Ley de Subvención Escolar Preferencial*. Diario Oficial de la República de Chile. <http://bcn.cl/2f6q4>
- Ley No. 20.370.* (2009, sep. 12). *Establece la Ley General de Educación*. Diario Oficial de la República de Chile.
<http://bcn.cl/2f73j>
- Ley No. 20.529.* (2011, ago. 27). *Sistema Nacional de Aseguramiento de la Calidad de la Educación Parvularia, Básica y Media y su Fiscalización*. Diario Oficial de la República de Chile.
<http://quality//bcn.cl/2oyqx>
- Maroy, C., & Voisin, A. (2015). Comparing accountability policy tools and rationales: various ways, various effects? In H. Kothhof, & E. Klerides (Eds.), *Governing educational spaces: Knowledge, teaching, and learning in transition* (pp. 35-58). Sense.
- Ministerio de Educación*. (2013a, oct. 25). *Decreto 381: Establece los Otros Indicadores de Calidad Educativa a que se refiere el artículo 3o, letra A), de la Ley No. 20.529, que establece el Sistema Nacional de Aseguramiento de la Calidad de la*

- Educación Parvularia, Básica y Media y su Fiscalización*. Diario Oficial de la República de Chile. <https://bit.ly/3BcEgYL>
- Ministerio de Educación*. (2013b). *Fundamentos Otros Indicadores de Calidad Educativa*. Unidad de Currículum y Evaluación. <https://bit.ly/2UNCn3X>
- Ministerio de Educación*. (2014b). *Otros Indicadores de Calidad Educativa*. Unidad de Currículum y Evaluación.
- Ministerio de Educación*. (MINEDUC, 2014, mar. 11). *Decreto 17: Aprueba la metodología de ordenación de todos los establecimientos educacionales reconocidos por el Estado, conforme a lo dispuesto en el inciso cuarto del artículo 17o. de la Ley No. 20.529*. Diario Oficial de la República de Chile.
- OCDE* (2004). *Evaluaciones de políticas nacionales de educación: Educación en Chile*. OCDE. <https://bit.ly/3xAN0Wd>
- OCDE*. (2017). *PISA 2015 results (Volume III): Students' well-being*. OCDE. <http://dx.doi.org/10.1787/9789264273856-en>
- OCDE* (2018). *El programa PISA de la OCDE: Qué es y para qué sirve*. OCDE. <https://bit.ly/3wFOVr9>
- Ortiz, I. (2012). En torno a la validez del sistema de medición de la calidad de la educación en Chile. *Estudios Pedagógicos*, 38(2), 355-373. <http://dx.doi.org/10.4067/S0718-07052012000200022>
- Ozga, J. (2014). Knowledge, inspection and the work of governing. *Journal of Education*, 2(1), 6-38. <https://doi.org/10.25749/sis.3876>
- Palestro, S. (2016). Androcentrismo en los textos escolares. En S. del Valle (Ed.), *Educación no sexista. hacia una real transformación* (15-23). Red Chilena contra la Violencia hacia las Mujeres.
- Parcerisa, L., & Villalobos, C. (2020). Movimientos sociales y resistencia al accountability en Chile: Estrategias discursivas, identidad y acciones de la campaña Alto al SIMCE. *Izquierdas*, 49, 2427-2455. http://www.izquierdas.cl/images/pdf/2020/n49/art115_2427_2455.pdf
- Pino, M., Oyarzún, G., & Salinas, I. (2016). Crítica a la rendición de cuentas: narrativa de resistencia al sistema de evaluación en Chile. *Cadernos CEDES*, 36(100), 337-354. <http://dx.doi.org/10.1590/cc0101-32622016171362>
- Piñuel, J. L. (2002). Epistemología, metodología y técnicas del análisis de contenido. *Estudios de Sociolingüística*, 3(1), 1-42. <https://journals.equinoxpub.com/SS/article/view/2459>
- PNUD, & UNICEF*. (2014). *El papel de la educación en la formación del bienestar subjetivo para el desarrollo humano: Una revisión al caso chileno*. UNICEF. <https://uni.cf/3kkucqg>
- PNUD*. (2012). *Informe de Desarrollo Humano en Chile. Bienestar Subjetivo: El desafío de repensar el desarrollo*. PNUD. <https://bit.ly/2U4RUfM>
- Prieto, A. (1983). *La modernización educacional*. Universidad Católica.
- Priori, L. (2003). Repositioning documents in social research. *Sociology*, 42(5), 821-836. <https://doi.org/10.1177/0038038508094564>
- Ramos, C. (2021). *Los directivos de establecimientos educacionales evalúan las evaluaciones: Reporte de una encuesta nacional*. Universidad Alberto Hurtado.
- Rosenkvist, M. A. (2010). *Using student test results for accountability and improvement: A literature review*. OECD Education Working Paper No. 54. OECD. <https://doi.org/10.1787/19939019>
- Torres, J., Contreras, S., Lippi, L., Huaiquimilla, M., & Leal, R. (2019). Hábitos de vida saludable como indicador de desarrollo personal y social: Discursos y prácticas en escuelas. *Calidad en la Educación*, 50, 357-392. <http://dx.doi.org/10.31619/caledu.n50.728>
- UNESCO*. (2005). *Informe de seguimiento de la EPT en el mundo 2005: Educación para todos. El Imperativo de la calidad*. <https://unesdoc.unesco.org/ark:/48223/pf0000150169>
- UNESCO*. (2009). *Sistema de indicadores para el seguimiento de las metas educativas de las cumbres de las Américas. Metodología de construcción y uso*. Proyecto Regional de Indicadores Educativos (PRIE). <https://bit.ly/36EKfxH>
- UNESCO*. (2017). *Informe de seguimiento de la educación en el mundo: Rendir cuentas en el ámbito de la educación; cumplir nuestros compromisos*. UNESCO. <https://bit.ly/3wHhDrA>
- UNICEF*. (2000). *Defining quality in education*. <https://bit.ly/3xSoDmQ>
- Verger, A., & Parcerisa, L. (2017). *Accountability and Education in the post-2015 scenario: International trends, enactment dynamics and socio-educational effects*. UNESCO.
- Verger, A., Parcerisa, L., & 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. <http://doi.org/10.1080/00131911.2019.1522045>
- World Bank*. (2007). *Chile institutional design for an effective education quality assurance*. <https://bit.ly/3yWc6Pr>

Sobre la autora y el autor:

Gonzalo Oyarzún Vargas es Profesor de Inglés, Mg. en Política Educativa (Univ. Alberto Hurtado), y académico de la Universidad Finis Terrae (Chile). Realiza docencia en la enseñanza del inglés y formación pedagógica en la educación superior. Participa en investigaciones relacionadas con políticas de mercado, rendición de cuentas y evaluación escolar en Chile.

Alejandra Falabella es Dra. en Sociología de la Educación (Univ. of London, UK), Mg. en Antropología Social (Univ. de Chile) y Educadora de Párvulos (Pontificia Universidad Católica de Chile). Académica del Depto. de Política Educativa y Desarrollo Escolar, Univ. Alberto Hurtado. Investiga sobre políticas de estandarización, evaluación y rendición de cuentas; mercados escolares, gestión y cultura escolar, y privatización; y políticas de educación inicial a nivel nacional e internacional.