

LÍMITES DE LAS EVALUACIONES ESTANDARIZADAS PARA COMPRENDER LOS DESEMPEÑOS ESCOLARES. EL CASO DE LOS OPERATIVOS *APRENDER* EN ARGENTINA

Gonzalo Gutiérrez

*El estudio no se mide por el número de
páginas leídas en una noche ni por la
cantidad de libros leídos en un semestre.
Estudiar no es un acto de consumir ideas
sino de crearlas y recrearlas.*

Paulo Freire

§1. Introducción

La evaluación educativa es uno de los pilares constitutivos del trabajo de enseñar que contribuye a mirar aprendizajes y experiencias de niños, niñas y jóvenes, así como a valorar sus logros y obstáculos en la apropiación de saberes escolares socialmente relevantes. Por ello, la preocupación del Estado por conocer los aprendizajes construidos por los/as estudiantes siempre representa una responsabilidad y oportunidad para apreciar en qué medida se atiende su derecho a acceder a los saberes culturales necesarios para su plena inserción ciudadana. En este sentido, las políticas de evaluación deberían interrogarse sobre un principio básico de justicia: ¿todos los niños, niñas y jóvenes han podido apropiarse de las experiencias y saberes culturalmente valiosos a los que tienen derecho y la escuela pone a su disposición? Atender esta pregunta supone resolver otra que la acompaña en forma tácita referida a la discusión y definición de los dispositivos e instrumentos de evaluación en la escala de los sistemas educativos que pueden brindar información relevante al respecto.

En este artículo analizaremos las limitaciones que posee el uso de la información proporcionada por las evaluaciones estandarizadas para identificar, comprender y reflejar la variedad y complejidad de aprendizajes y experiencias formativas que los estudiantes construyen en la escuela. Para ello, se analizarán algunos problemas de validez que suelen presentarse en las comparaciones entre desempeños escolares de sistemas educativos y contextos sociales muy diferentes al de nuestro país. Ambas cuestiones dan sustento

al argumento según el cual no es posible reducir las interpretaciones sobre la calidad de los sistemas educativos a los resultados de las evaluaciones estandarizadas. Finalmente se presentarán algunas características y debilidades políticas, técnicas y pedagógicas que posee el Operativo *Aprender* en Argentina, las discusiones generadas sobre sus resultados y los desafíos pendientes para desarrollar un sistema democrático e integral de evaluación que se configure como insumo de relevancia para el desarrollo de políticas de enseñanza.

§2. Sobran evaluaciones, faltan consensos sobre su función y legitimidad político-pedagógica

Un debate importante viene produciéndose en nuestro país sobre los desempeños escolares, la calidad educativa y los instrumentos que posibilitan obtener información relevante para la toma de decisiones desde las políticas educativas. En este sentido, las pruebas estandarizadas vienen ocupando un lugar creciente en las referencias para analizar los logros y obstáculos del sistema educativo argentino.

Una perspectiva histórica de este asunto permite apreciar que en los últimos 24 años se ha incrementado en forma constante la aplicación de evaluaciones estandarizadas. Entre ellas se destacan el Programa para la Evaluación Internacional de Alumnos (PISA, por sus siglas en inglés) que depende de la Organización para la Cooperación y el Desarrollo Económicos (OCDE), el Tercer Estudio Evaluativo y Comparativo (TERCE) desarrollado desde UNESCO y destinado a estudiantes de nivel primario, así como el Estudio de Tendencias en Matemáticas y Ciencias (TIMSS, por sus siglas en inglés) y el Estudio Internacional de Progreso en Comprensión Lectora. Entre las evaluaciones nacionales, desde el año 1993, se encuentran los Operativos Nacionales de Evaluación (ONE) que desde 2016, fueron reemplazados por los *Aprender* y, desde el año 2017 se implementa un nuevo operativo de evaluación, denominado *Enseñar* dirigido a estudiantes del último año de los profesorados de Educación Primaria y Secundaria (Matemática, Lengua/ Lengua y Literatura, Historia, Geografía, Biología, Física, Química e Inglés). También se han desarrollado experiencias provinciales de evaluación en todo el país, por ejemplo, en Córdoba, cuatro operativos de evaluación, tres de ellos entre 1992 y 1994, el restante en 2005.

Las actuales autoridades educativas nacionales vienen planteando la necesidad de construir una cultura evaluativa que permita conocer los aprendizajes de los/as estudiantes, sin advertir que en los hechos, tal cultura ya se encuentra instalada aunque con escasas articulaciones entre el uso de la información brindada por las evaluaciones y el desarrollo de políticas de enseñanza. Por ejemplo, en la provincia de Córdoba, entre los años 1993 y 2017, se desarrollaron por lo menos, 37 evaluaciones estandarizadas a las que se deben añadir las correspondientes a programas nacionales y provinciales implementados en las escuelas que, según Gutierrez (2011), en el año 2010, llegaban a 58. Junto a ellas se encuentran también, las realizadas al interior de las escuelas en el marco de proyectos institucionales y/o curriculares de carácter más específico. Es decir, si sobre algo se ha construido experiencia en nuestro país, es sobre el desarrollo de modelos de evaluación educativa en diferentes escalas atendiendo a múltiples finalidades. En este sentido, es posible sostener la presencia de múltiples dispositivos de este tipo, que desplazan la centralidad del

trabajo escolar: enseñar, generando una saturación de información que reduce su relevancia y complejiza el reconocimiento de tendencias en los logros y obstáculos que sobre el desempeño escolar se vienen configurando en nuestro país. Como veremos más adelante, esta situación se refleja en el uso descontextualizado de medidas de tendencia central que al presentar indicadores en porcentajes sin el análisis de su evolución histórica, ni contar con recaudos metodológicos en las comparaciones de rendimientos entre poblaciones estudiantiles con condiciones materiales de vida y de escolarización muy diferentes entre sí, producen interpretaciones con problemas de validez en algunos casos y/o sesgadas en otros, que reflejan el estado de situación existente en torno a los desempeños escolares.

La multiplicidad y periodicidad de evaluaciones estandarizadas no supone ni implica consenso sobre los alcances que poseen sus conclusiones, su validez metodológica y ni sus implicancias para la construcción de políticas públicas. Por ello, gran parte de los conflictos en torno a ellas se debe a que el interés por conocer el lugar que ocupa nuestro país en el ranking de participantes o las comparaciones de desempeño entre provincias, suele articularse con la intencionalidad de cuestionar la relevancia de la educación pública poniendo bajo sospecha el trabajo docente y cuestionando la democratización producida durante las últimas décadas en nuestro país con las políticas de inclusión educativa.

§3. Los límites de las evaluaciones estandarizadas

La publicación de resultados sobre evaluaciones estandarizadas suele generar confusiones importantes porque no pueden, ni es su propósito, reflejar todo lo que se espera de la escuela ni lo que cotidianamente en ella se realiza. Usualmente, ellas no pueden dar cuenta de la presencia de experiencias participativas; ni de la construcción de confianza y dignidad que el trabajo docente posibilita a diario con miles de estudiantes y sus familias, ni de los diferentes modos de relación con el saber y la cultura generados por medio de la literatura, el arte y los múltiples recursos tecnológicos que se encuentran disponibles en las escuelas. Tampoco de la manera en que los niños/as y jóvenes construyen, en diálogo con el trabajo escolar, prácticas de ciudadanía significativas; y son justamente estas cuestiones, que también forman parte de los objetivos centrales del curriculum escolar y constituyen el núcleo que otorga identidad a la escuela, las que no logran ser reconocidas ni visibilizadas en las evaluaciones estandarizadas.

La mayoría de las evaluaciones estandarizadas toman como referencia cuatro áreas de conocimiento: Lengua (o lectura y escritura), Ciencias Naturales (o simplemente Ciencias), Matemática y Ciencias Sociales. Aquí opera una reducción del curriculum a lo evaluable que actúa como legitimante de lo enseñado, pues en la escuela se enseñan muchos más objetos y áreas de conocimiento que varían en su ubicación al interior de la trayectoria escolar teórica de cada sistema o sub-sistema educativo, por lo que en numerosas ocasiones, se construyen interpretaciones sobre desempeños escolares y calidad educativa a partir de dispositivos que evalúan cuestiones no necesariamente enseñadas por igual y/o en el mismo tramo escolar, en cada región, provincia y/o país. Así por ejemplo, PISA toma como referencia a jóvenes de 15 años sin considerar si asisten a la escuela secundaria, el año en que lo hacen e incluso si se encuentran en el nivel primario, por lo que sus resultados permiten hablar de los desempeños escolares de un grupo etario entre los países

participantes, pero no de la calidad de sus sistemas educativos. Una segunda reducción se produce de la evaluación en su sentido más amplio, a la medición de indicadores limitados para comunicar la variedad de aprendizajes y experiencias construidas por los estudiantes. Este asunto adquiere relevancia, porque numerosos desempeños escolares no pueden conocerse por medio de las evaluaciones estandarizadas y requieren de otros tipos de instrumentos. Así por ejemplo, es creciente el número de escuelas que anualmente participan en Ferias de Ciencia nacionales y provinciales; también se incrementan las distinciones y reconocimientos de estudiantes argentinos en diferentes exposiciones internacionales, tanto por sus logros académicos, como por sus innovaciones tecnológicas; es cada vez más frecuente el desarrollo de proyectos escolares que promoviendo el desarrollo de prácticas solidarias, atienden como parte de sus procesos formativos diferentes necesidades de sus comunidades¹. Las evaluaciones estandarizadas tampoco pueden dar respuestas a interrogantes nodales para comprender la calidad de un sistema educativo, tales como los siguientes: ¿cómo se mide el cuidado de la infancia y la adolescencia realizado por los docentes? ¿Cómo se miden las experiencias significativas que posibilitan a los niños/as y jóvenes sentirse parte de la escuela, la sociedad y el mundo? ¿Cómo se miden las relaciones pedagógicas que han posibilitado en estos años a miles de estudiantes volver a la escuela o mantenerse en ella, asumiéndose como sujetos de derecho, con plenas capacidades de aprender? Claramente no son este tipo de evaluaciones el mejor modo de responder estas preguntas y sus resultados en forma aislada, no pueden reflejar la calidad de un sistema educativo. Por ello, es posible sostener que una política de evaluación sin perspectiva pedagógica, asentada sólo en la preocupación de medir y controlar, difícilmente pueda generar medidas que construyan mejores oportunidades de aprender para los niños, niñas y jóvenes.

§4. Problemas de validez en la comparación de los desempeños escolares entre países

Una de las cuestiones más problemáticas en algunos análisis de tendencias educativas, consiste en el uso fragmentario y descontextualizado de estadísticas que habilitan a realizar afirmaciones políticas y pedagógicas que carecen de sustento. En este sentido, es frecuente observar comparaciones lineales de logros y aprendizajes educativos entre países, sin considerar las diferentes tradiciones de escolarización existentes, sus puntos de partida y los modos de concebir la relación Estado-educación reflejados en políticas de financiamiento educativo, formación docente, provisión de recursos para la enseñanza y apoyo al trabajo de enseñar que en conjunto, configuran condiciones de enseñanza muy

¹ Así por ejemplo, en 2014, el Ministerio de Educación de la Nación reconoció con el premio presidencial Escuelas Solidarias a cuatro escuelas de Buenos Aires, Córdoba, La Pampa y Jujuy. En ese mismo año, durante la Feria Provincial de Ciencias realizada en la provincia de Córdoba, 34 proyectos escolares pasaron a la instancia nacional en Tecnópolis. Algunos de los proyectos que fueron seleccionados para participar de la instancia nacional, fueron los siguientes: IPET 44, de Miramar, con un trabajo sobre abono orgánico para el cultivo de trigo. Su propuesta consistía en usar desechos orgánicos para generar biogás y biolito, un abono particularmente provechoso para el trigo; IPET 85 junto con la escuela República de Italia (Estación General Paz) donde los/as estudiantes desarrollaron un proyecto sobre caramelos de goma de nopal (cactus del norte cordobés) logrando determinar que ese era el alimento nutritivo más aceptado por sus compañeros; El IPEM 274, de Huinca Renancó, indagó en el sexting (el envío por celular de fotos y mensajes con contenidos sexuales). Allí, estudiantes y docentes estudiaron el alcance que tiene esta práctica en las escuelas de la localidad y analizaron sus riesgos; El IPET 80 de Río Cuarto desarrolló un sistema automático que permite encender y apagar la iluminación de un espacio.

diferentes entre países y regiones. En este sentido, es posible apreciar que en los últimos años, numerosos estudios comparan la educación argentina con la de Chile y Finlandia omitiendo dar cuenta de las profundas diferencias entre nuestros sistemas y tradiciones escolares. Así, por ejemplo, si bien es cierto que Chile posee indicadores de desempeño escolar mejores que Argentina, también es verdad que cuenta con uno de los sistemas más desiguales del mundo, según un informe de la OCDE en 2004 (ver OCDE, 2015) y otro de UNESCO en 2013 (Muñoz, 2013). La segregación se produce en el vecino país por dos vías: el pago complementario que las familias realizan por su educación, de modo tal que a mayor ingreso, mayores posibilidades de pagar y obtener una mejor educación; y el hecho de que las escuelas pueden elegir a sus estudiantes. De este modo, quienes pertenecen a sectores de menores ingresos no pueden romper con los círculos de reproducción de la pobreza. Por su parte en Argentina, la educación pública (de gestión estatal y privada) es solventada en un 95 % por el Estado y solo el 5 % por las familias (González Olguín, 2015), además, la escuela pública estatal no puede generar políticas de selección de su matrícula, porque, es un pilar del derecho educativo, es para todas/os.

Las comparaciones con Finlandia, también omiten cuestiones relevantes para comprender las diferencias existentes entre su sistema educativo y el nuestro. En el país nórdico, el 92 % de su sistema es público, se comienza a ir a la escuela a los 7 años, no tiene pobreza infantil, hace más de 30 años universalizó el acceso a la escolaridad, su sistema educativo es menor al de la provincia de Buenos Aires, posee consensos de décadas sobre los horizontes de sus políticas públicas y el financiamiento educativo se logra con una presión tributaria del 43 % (año 2012, según OCDE). A diferencia de esta realidad, en nuestro país, la inversión educativa se logra con una presión tributaria del 29,5 % (año 2012, según OCDE); hace 15 años, la mitad de la población infantil estaba en condiciones de pobreza, en 30 años se ha duplicado el acceso a la secundaria pasando del 42 % al 85 % en 2013 (SITEAL) y las políticas públicas poseen una alta inestabilidad por las disputas entre posiciones que consideran al conocimiento como un bien de mercado, que debe regirse por la lógica de la oferta y la demanda y quienes lo entienden como un derecho y al Estado como responsable de generar las condiciones para que todos los niños/as y jóvenes puedan acceder a él, independientemente de su nivel socioeconómico, género o región en que se encuentre.

Con el desempeño escolar en Argentina también suelen realizarse comparaciones problemáticas. Así por ejemplo, es cierto que en las evaluaciones de PISA se pasó a ocupar uno de los últimos lugares entre los países participantes, pero también lo es que Argentina posee la tasa más alta de escolarización, por encima de Chile, Uruguay y Brasil (SITEAL, 2017). Es cierto que en los ONE 2013, entre un cuarto y un tercio de los estudiantes del último año de la secundaria tuvieron un desempeño bajo, pero también hay que señalar que dicho porcentaje fue menor en tres de las cuatro áreas evaluadas al de 2010. Lo mismo ocurrió con los datos arrojados por el *Aprender 2016* donde, como veremos más adelante, salvo en dos áreas de la escuela secundaria (Matemática y Ciencias Naturales), en el resto, el desempeño mejoró con respecto a 2013. Es a partir de estos datos, que sostenemos la imposibilidad de interpretar adecuadamente la situación de los desempeños escolares sin una perspectiva histórica que permita reconocer su tendencia. Algo similar ocurre cuando

se describe como una situación de crisis terminal el ciclo básico de la escuela secundaria donde se encuentran los desempeños escolares más bajos y los mayores porcentajes de repitencia y abandono escolar, sin considerar que es en ese tramo de escolaridad donde se ha producido la mayor incorporación de jóvenes que constituyen la primera generación familiar en acceder a dicho nivel educativo y por tanto, donde se concentran los desafíos más complejos de resolver pedagógicamente (Gutierrez et al., 2014). Estas diferencias en los criterios de interpretación muestran la relevancia de analizar el modo en que los datos se utilizan, porque de allí se ponen en juego cuestiones nodales sobre la validez política, pedagógica y metodológica de las interpretaciones que sobre ellos se realizan.

§5. El Operativo *Aprender*: características y limitaciones de un nuevo dispositivo de evaluación estandarizado

En noviembre de 2017 el Ministerio de Educación Nacional implementó dos operativos nacionales de evaluación. Uno dirigido a quienes cursan el nivel primario y secundario; y otro a estudiantes del último año de los profesorados de Educación Primaria y de Educación Secundaria (Matemática, Lengua/ Lengua y Literatura, Historia, Geografía, Biología, Física, Química e Inglés). *Aprender* en un caso y *Enseñar* en el otro, se configuran como nuevos mecanismos para evaluar desempeños estudiantiles con fuertes cuestionamientos públicos por sus problemas de validez y los débiles consensos académicos y sociales con los que se han construido.

Tanto *Aprender* como *Enseñar*, se inscriben en lo previsto por la Ley de Educación Nacional n°26.206 donde se sostiene que la evaluación de los aprendizajes es una obligación del Estado prevista en el artículo 85 inciso d), según el cual el Estado debe implementar "...una política de evaluación concebida como instrumento de mejora de la calidad de la educación" con la finalidad de "... garantizar las condiciones materiales y culturales para que todos/as los/as alumnos/as logren aprendizajes comunes de buena calidad, independientemente de su origen social, radicación geográfica, género o identidad cultural". Según lo establecido en el artículo 94 de dicha ley, es responsabilidad del Estado a través de su Ministerio de Educación, desarrollar e implementar "...una política de información y evaluación continua y periódica del sistema educativo para la toma de decisiones tendiente al mejoramiento de la calidad de la educación, la justicia social en la asignación de recursos, la transparencia y la participación social". La Ley de Educación Nacional plantea una perspectiva amplia sobre la evaluación al señalar en su artículo 95 que:

Son objeto de información y evaluación las principales variables de funcionamiento del sistema, tales como cobertura, repetición, deserción, egreso, promoción, sobre-edad, origen socioeconómico, inversiones y costos, los procesos y logros de aprendizaje, los proyectos y programas educativos, la formación y las prácticas de docentes, directivos y supervisores, las unidades escolares, los contextos socioculturales del aprendizaje y los propios métodos de evaluación.

Es decir, las evaluaciones de desempeños en Argentina se encuentran inscriptas en regulaciones según las cuales la calidad educativa no puede reducirse a lo medido por

pruebas estandarizadas y sus resultados deben analizarse en relación a otras variables como terminalidad, repitencia, sobreedad, etc.

En tanto la educación es un bien público, las decisiones sobre finalidades, modalidades y destinatarios de la evaluación que pretenda conocer el estado de la calidad educativa, no pueden construirse unilateralmente y requieren de la participación de múltiples actores institucionales, tales como universidades públicas y privadas, representantes de la sociedad civil, organizaciones sindicales docentes, investigadores. Así lo plantea el artículo 98 de la Ley de Educación Nacional donde se señala:

Créase el Consejo Nacional de Calidad de la Educación, en el ámbito del Ministerio de Educación, Ciencia y Tecnología, como órgano de asesoramiento especializado, que estará integrado por miembros de la comunidad académica y científica de reconocida trayectoria en la materia, representantes de dicho Ministerio, del Consejo Federal de Educación, del Congreso Nacional, de las organizaciones del trabajo y la producción, y de las organizaciones gremiales docentes con personería nacional.

Las funciones principales de este Consejo Nacional de Calidad son:

a) Proponer criterios y modalidades en los procesos evaluativos del Sistema Educativo Nacional. b) Participar en el seguimiento de los procesos de evaluación del Sistema Educativo Nacional, y emitir opinión técnica al respecto. c) Elevar al Ministerio de Educación, Ciencia y Tecnología propuestas y estudios destinados a mejorar la calidad de la educación nacional y la equidad en la asignación de recursos. d) Participar en la difusión y utilización de la información generada por dichos procesos. e) Asesorar al Ministerio de Educación, Ciencia y Tecnología con respecto a la participación en operativos internacionales de evaluación.

En este marco, resulta problemático el bajo consenso que poseen *Aprender* y *Enseñar* elaborados en soledad por las autoridades nacionales de educación.

En 2017, el Operativo *Aprender* continúa evaluando a estudiantes de nivel primario y secundario, aunque con algunas diferencias con respecto a su versión 2016 (Cuadro 1) por las áreas y grados considerados en nivel primario. Una cuestión novedosa en los cambios introducidos, consiste en la evaluación a estudiantes de 4^{to} grado de primaria en producción lectora. Por su parte el Operativo *Enseñar* se presenta como algo diferente y nuevo con respecto a las tres evaluaciones sobre diseño curricular implementadas entre los años 2011-2015. A continuación, nos centraremos en algunas características y limitaciones que poseen las evaluaciones estandarizadas denominadas *Aprender*.

Operativo de evaluación	Niveles y grados/años	Tipo de muestra	Áreas evaluadas
<i>Aprender 2016</i>	3 ^{er} grado Primaria	Muestral	Lengua y Matemática
	6 ^{to} grado Primaria	Censal	
	2 ^{do} y 3 ^{er} años Secundaria	Muestral	
	5 ^{to} y 6 ^{to} años Secundaria	Censal	Lengua, Cs. Sociales y Cs. Naturales
<i>Aprender 2017</i>	4 ^{to} grado Primaria	Muestral	Producción escrita
	6 ^{to} grado Primaria	Censal	Ciencias Sociales y Ciencias Naturales
	6 ^{to} años Secundaria	Censal	Matemática y Lengua
<i>Enseñar 2017</i>	Estudiantes del último año de institutos de formación docente que estén realizando su residencia en profesorados de Educación Primaria y de profesorados de materias del ciclo básico de Educación Secundaria (Matemática, Lengua / Lengua y Literatura, Historia, Geografía, Biología, Física, Química e Inglés).	Muestra representativa (público-privada). Algunas jurisdicciones eligieron realizar una muestra censal, otras muestral. Quedan excluidos institutos con menos de 15 estudiantes en 4 ^{to} año.	Comunicación escrita (lectura y escritura) y criterio pedagógico (planificación de la enseñanza, implementación de estrategias de enseñanza y evaluación de aprendizajes).

CUADRO 1. Operativos *Aprender* 2016-2017 y *Enseñar* 2017 según destinatarios, tipo de muestra y área evaluada

§6. Las limitaciones del Operativo *Aprender* como insumo para conocer los aprendizajes estudiantiles

Las evaluaciones denominadas *Aprender* han tenido numerosos cuestionamientos por problemas metodológicos, perspectivas pedagógicas e intencionalidades políticas. En documentos oficiales se sostiene que aunque las recomendaciones internacionales en torno a evaluaciones estandarizadas sugieren una periodicidad de dos o tres años que eviten la saturación de información, se ha optado por su aplicación anual. Su fundamento es la "...necesidad de recrear una dinámica que nos vuelva a conciliar con instancias evaluativas..." (Documento Marco *Aprender* 2016). De este modo, el control y la medición de resultados toman forma como marcas identitarias de este nuevo dispositivo de evaluación.

Metodológicamente, se observa un cambio significativo entre los ONE y *Aprender*. En el primero de ellos, los desempeños se clasificaban en torno a tres categorías: Alto, Medio y Bajo, mientras que en *Aprender*, se lo hace en cuatro: Avanzado, Satisfactorio, Básico y

por debajo del Básico. En este cambio, cada categoría informa sobre niveles de desempeño diferentes, no sobre saberes, ello hace que no sea válido establecer agrupamientos entre ellas para dar cuenta de tendencias en el desempeño. Este nuevo modo de informar la producción de los/as estudiantes da cuenta de un problema conceptual aún no despejado metodológicamente: si las categorías informan sobre desempeños, es decir, lo que pueden o no hacer los/as estudiantes, ello no implica que lo hagan sobre sus aprendizajes, porque el saber no necesariamente se reduce al hacer. Esta ambigüedad se expresa en el corte establecidos entre categorías, donde el Básico se define como una proporción del 25 % por debajo de la categoría Satisfactorio sin que pueda precisarse a qué tipo de desempeño ni saberes remite esta diferencia. El análisis de *Aprender* como instrumento de evaluación permite apreciar limitaciones similares a las que en otros tiempos tuvieron los ONE, al dar cuenta de los aprendizajes estudiantiles a partir de la medición de logros en solo cuatro áreas del curriculum escolar, sin considerar otros campos de conocimiento, ni indicadores de desempeño alternativos producidos en el marco de la participación creciente de escuelas en Ferias de Ciencias nacionales y provinciales; el desarrollo de proyectos innovadores de escuelas técnicas; experiencias de trabajo comunitario realizados a lo largo y ancho del país por parte de miles de niños y niñas con sus docentes. Se invisibilizan así, aprendizajes valiosos construidos en espacios como jornada extendida, talleres de escuelas técnicas, Centros de Actividades Juveniles o programas provinciales tales como el Programa de Inclusión y Terminalidad donde el gobierno de la provincia de Córdoba ha destinado importantes esfuerzos para la reinserción educativa de miles de jóvenes.

Como instrumento de evaluación, *Aprender* invisibiliza dimensiones centrales del aprendizaje y planos del trabajo de enseñar sostenidos por los docentes en función de demandas específicas que al respecto construye la misma política pública por medio de los ministerios provinciales de educación. Una mirada comparativa con respecto a los ONE donde se aplicaban evaluaciones en 3^{er} grado y 6^{to} grado de la escuela primaria y 2^{do}/3^{er} año y 5^{to}/6^{to} año de secundaria, muestra que en *Aprender* 2017 se reducen los años a evaluar, al no incluir 3^{er} grado de primaria, ni 2^{do}/3^{er} año de secundaria, brindando de este modo, menos información sobre los aprendizajes estudiantiles en el punto más crítico de la escolaridad secundaria (primer ciclo) que es donde se ubican las mayores dificultades y problemáticas en la enseñanza y convivencia escolar. Pero además, en primaria, *Aprender* 2017, alterna las áreas evaluadas incluyendo Ciencias Sociales y Naturales y omitiendo Matemática y Lengua, sin que queden claros los criterios metodológicos en que se funda dicha decisión, aunque todo hace pensar que se evaluarán en adelante de forma bianual.

Por otro lado, la eliminación de ítems abiertos en *Aprender*, con respecto a los ONE, genera otro tipo de reducción pedagógica, al excluir de la evaluación contenidos y competencias vinculados con las posibilidades de describir, explicar y argumentar. Esta opción por trabajar solo con ítems cerrados (múltiple opción) representa un retroceso en la calidad de la información que puede brindar sobre los aprendizajes. Las respuestas de ítem cerrado no permiten conocer si las respuestas adecuadas corresponden a comprensiones genuinas de los estudiantes o al azar por haber acertado en la opción elegida. Pero tampoco permiten generar hipótesis sobre los modos de razonamiento, sus características y las dificultades que presentan para resolver adecuadamente las opciones presentadas.

De este modo, los resultados del *Aprender* pueden informar sobre cuántos estudiantes contestaron bien o mal las preguntas planteadas, pero no sobre los aprendizajes que han construido.

La opción de presentar solo ítems cerrados de múltiple opción representa un retroceso en las formas de construir información sobre los aprendizajes logrados por niños, niñas y jóvenes. Durante los últimos años, escuelas y docentes vienen desarrollando proyectos pedagógicos promovidos por el propio Estado como políticas curriculares y de enseñanza que posibiliten a los estudiantes "leer el mundo y transformarlo" por medio de descripciones, interpretaciones, argumentaciones, explicaciones, desarrollando hipótesis sobre diferentes aspectos de la relación con el saber. Sin embargo, las evaluaciones *Aprender* privilegian un modo de evaluar que excluye la consideración sobre estos aprendizajes, que es donde mayor esfuerzo se viene realizando desde el trabajo de enseñar. De este modo, al no preverse preguntas abiertas se pierde la posibilidad de conocer las fortalezas y debilidades que presentan los aprendizajes estudiantiles, pero también se reducen las posibilidades de elaborar a partir de los resultados obtenidos, orientaciones para la enseñanza y aportes para el desarrollo de políticas de formación docente continua.

Las evaluaciones *Aprender* comienzan a condicionar la enseñanza, al hacer que poco a poco, esta empiece a organizarse por lo privilegiado en las evaluaciones y no por la relevancia curricular de aquello que se decide transmitir en la escuela. En Córdoba y en el país, por "recomendaciones de las autoridades", cada vez más escuelas destinan tiempo de enseñanza a preparar a sus estudiantes en los modos de responder las actividades de simulación publicadas en páginas oficiales del Ministerio de Educación y Deportes de la Nación. Se condiciona de ese modo las respuestas que pasan a ser el resultado de un "entrenamiento" previo y la enseñanza deja de tener como horizonte, la distribución de bienes culturales para comenzar a centrarse en que los resultados de las evaluaciones "den bien". Este efecto se condice con lo planteado en documentos que señalan "...la regularidad del operativo permitirá incorporar la evaluación en el calendario escolar, planificar acciones basadas en los resultados de operativos y fomentar una cultura de difusión y uso de información para la mejora" (Documento Marco Aprender 2016). De este modo, la enseñanza corre el riesgo de quedar subordinada a la lógica de la evaluación, que deja de ser pensada como insumo para conocer y comprender los aprendizajes alcanzados por los niños, niñas y jóvenes y se transforma en un dispositivo de control, medición y comparación de resultados entre estudiantes, escuelas, provincias y regiones, que dan lugar de hecho (aunque no se lo enuncie de ese modo), a la elaboración de rankings, donde unos/as están mejor que otros. Complementariamente, este modo de asumir la evaluación, tiende a instrumentalizar y despolitizar el sentido del trabajo de enseñar, de modo tal que los docentes comienzan a ser concebidos como aplicadores de instrumentos diseñados externamente.

Finalmente, podemos señalar que en las evaluaciones *Aprender*, la meritocracia se hace lugar: Se sostiene en los documentos oficiales que: "En el caso de los estudiantes de 5^{to}/6^{to} que desean estudiar carreras vinculadas a la docencia, se les dará la opción de dejar sus datos personales para participar de un proceso de selección para un programa de becas". Es decir, la historia del desempeño escolar comienza a funcionar estigmatizantemente,

porque poco a poco la escuela corre el riesgo de transformar las diferencias de rendimiento académico en diferencias de oportunidades sociales.

§7. Los resultados del Aprender 2016 en debate

Las controversias en torno al *Aprender* se trasladaron a la lectura de sus resultados. En marzo de 2017, en medio de un conflicto nacional con los gremios docentes por la no convocatoria a paritarias docentes previstas por la Ley de Educación Nacional, las autoridades nacionales presentaron los resultados del operativo de evaluación implementado en octubre de 2016. En su interpretación de los datos obtenidos, se planteó un escenario de crisis educativa al sostener que casi el 70 % de los estudiantes de nivel secundario no podían realizar operaciones básicas de Matemática y que algo similar ocurría en Ciencias Naturales al agrupar los desempeños "por debajo de Básico y Satisfactorio" que ya hemos señalado no es válido metodológicamente. Pero además, se realizó una comparación entre *Aprender* y ONE 2013, sin compatibilizar las categorías utilizadas en ambas evaluaciones. Cuando meses más tarde se realizaron los procedimientos técnicos que permiten comparar los resultados de ambas evaluaciones, pudo verse que la situación educativa es muy diferente a la anunciada a comienzos del ciclo escolar.

Es verdad que los desempeños escolares deben mejorar y están lejos de lo que quisiéramos como sociedad. Pero también, que la comparación del Operativo *Aprender* 2016, con el ONE 2013, refleja en términos generales, una mejora de los desempeños escolares. Según el "Informe de resultados *Aprender* 2016", en nivel primario (Cuadro 2) puede apreciarse que tanto en 3^{er} grado como en 6^{to} grado, en Lengua y Matemática, el porcentaje de estudiantes con niveles de logro Avanzado y Satisfactorio creció entre ONE 2013 y *Aprender* 2016 y se redujo el de estudiantes en "Por debajo del nivel básico".

Desempeño	Primaria 3 ^{er} grado				Primaria 6 ^{to} grado			
	Lengua		Matemática		Lengua		Matemática	
	2013	2016	2013	2016	2013	2016	2013	2016
Avanzado	34	41	12	12	23,6	32,3	12,3	19,7
Satisfactorio	25	21	47	51	35	34,5	39,4	38,9
Básico	26	22	18	22	23,7	18,7	26,5	23,4
Por debajo del nivel básico	14	16	22	15	18	14,7	21,8	18

CUADRO 2. Comparación de desempeños entre el Operativo Nacional de Evaluación 2013 y el *Aprender* 2016 para nivel secundario, en porcentajes. **Fuente:** Elaboración propia en base a datos publicados en *Aprender* (2016)

En el nivel secundario existe una mayor complejidad pedagógica expresada tanto en el crecimiento de su cobertura en las últimas décadas, como en el hecho de que una proporción importante de quienes acceden a este nivel educativo constituyen la primera generación familiar en lograrlo. Esta mayor complejidad también se refleja en los desempeños escolares. Según el informe de resultados de *Aprender* 2016 en secundaria (Cuadro 3), para 2^{do}/3^{er} año área Lengua, puede apreciarse que el porcentaje de estudiantes con niveles de logro Avanzado y Satisfactorio subió levemente entre ONE 2013 y *Aprender* 2016, aunque también lo hizo el porcentaje de "Por debajo del nivel básico". En Matemática se mantuvo el porcentaje de estudiantes con niveles de logro Avanzado y Satisfactorio y se incrementó el de estudiantes en "Por debajo del nivel básico". En 5^{to}/6^{to} año Secundaria las áreas Lengua y Ciencias Sociales, muestran que el porcentaje de estudiantes con niveles de logro Avanzado y Satisfactorio subió levemente entre ONE 2013 y *Aprender* 2016 y se redujo levemente el de "Por debajo del nivel básico". En cambio en Ciencias Naturales aunque creció levemente el porcentaje de estudiantes con niveles de logro Avanzado y Satisfactorio, también se incrementó el porcentaje de estudiantes en "Por debajo del nivel básico". Es Matemática el área más crítica entre las evaluadas en tanto allí el porcentaje de estudiantes con niveles de logro Avanzado y Satisfactorio disminuyó entre ONE 2013 y *Aprender* 2016 y se incrementó levemente el porcentaje de estudiantes "Por debajo del nivel básico".

Desempeño	Secundaria 2 ^{do} /3 ^{er} año				Secundaria 5 ^{to} /6 ^{to} año							
	Lengua		Matemática		Lengua		Matemática		C. Naturales		C. Sociales	
	2013	2016	2013	2016	2013	2016	2013	2016	2013	2016	2013	2016
Avanzado	23	27	5	9	11	9,4	7,4	5,2	3,1	10,1	32,8	34
Satisfactorio	37	35	33	29	39,5	44,2	27,9	24,6	57,2	53,6	23,5	24,9
Básico	26	21	36	28	21	23,4	24,7	29,3	23,8	19,4	24	22,3
Por debajo del Nivel Básico	14	17	27	34	28,5	23	40	40,9	15,9	16,9	19,1	18,8

CUADRO 3. Comparación de desempeños entre el Operativo Nacional de Evaluación 2013 y el *Aprender* 2016 para nivel secundario, en porcentajes. Fuente: Elaboración propia en base a datos publicados en *Aprender* (2016)

El análisis de la comparación entre ONE 2013-*Aprender* 2016, refuerza lo sostenido en un documento de reciente circulación elaborado por CTERA (2017, Los operativos de evaluación "*Aprender* y Enseñar"), donde se sostiene que la comparación entre el ONE 2013 y *Aprender* 2016 corrobora lo que otras evaluaciones (TERCE y PISA, entre ellas) venían mostrando: el nivel educativo en Argentina no ha bajado. Al contrario, hay evidencias de mejora, especialmente en la educación primaria y en secundaria, si bien la mejora se evidencia en algunas áreas (Lengua y Sociales), en Matemática y Naturales se presenta una situación de estancamiento en cuanto a los estudiantes que no alcanzan el nivel básico,

ya que estos porcentajes descienden muy levemente. Por ello, es necesario preguntarse en qué medida, la interpretación sesgada de la información surgida del *Aprender* 2016 busca construir una sensación de "catástrofe educativa" que no se deriva de los desempeños estudiantiles. Algo similar ocurre con la comparación de resultados entre quienes asisten a escuelas de gestión estatal y privada. Mientras el gobierno señala que existe una brecha abismal entre ambas, estudios recientes de CTERA (2017) mostraron que al controlar las variables *nivel socioeconómico y disponibilidad de recursos didácticos en las escuelas*, dichas diferencias desaparecen e incluso se invierten, por lo que no es posible sostener que la educación brindada en el sector privado es mejor que la ofrecida en el estatal.

El análisis realizado sobre las evaluaciones *Aprender*, permite sostener que se han construido con escasos consensos político-académicos y que el instrumento implementado tiene numerosos problemas técnico-pedagógicos, porque evalúa lo que la escuela no enseña y de un modo en que la escuela tampoco lo hace. Frente a esta situación cabe preguntarse ¿Qué enseña la escuela? y ¿Si los *Aprender* evalúan en forma diferente a como lo hacen los docentes, en qué radica dicha diferencia? En primer lugar habría que señalar que en las escuelas se enseña aquello que, prescripto en los diseños curriculares, es definido como prioritario por cada una de las jurisdicciones de un modo muy diferente al propuesto por las evaluaciones *Aprender*. Mientras las escuelas enseñan a resolver situaciones problemáticas, las evaluaciones *Aprender* presentan pseudo-problemas; mientras los docentes enseñan múltiples modos de argumentar las interpretaciones, explicaciones y sensaciones, *Aprender* presenta actividades cerradas de múltiple opción. Mientras a las escuelas y sus docentes se les demanda promover en sus estudiantes la revisión y reflexión permanente de sus producciones y procesos de aprendizaje, en las evaluaciones *Aprender*, se presentan entre 25/30 actividades para responder en una hora, a razón de dos minutos por pregunta.

§8. Por políticas integrales de evaluación

En la actualidad, a diferencia de décadas anteriores, para nuestra sociedad no resulta indiferente si se aprende o no en la escuela, si los/as estudiantes la concluyen en los tiempos previstos y si logran continuar estudios superiores. Estos desafíos se inscriben en esfuerzos más amplios por construir una sociedad más justa. Por ello, sostenemos que el interés por conocer los aprendizajes logrados por los niños, niñas y jóvenes, es un derecho social y la responsabilidad del Estado es construir dispositivos adecuados para ello. Para ello, resulta central desarrollar políticas integrales y participativas de evaluación que generen información válida y relevante para elaborar dispositivos de formación docente y estructuras de apoyo al trabajo de enseñar que permitan sostener los logros alcanzados y revertir las dificultades encontradas. Esto requiere políticas de evaluación situadas, permanentes, colectivas, formativas, es decir, integrales y no punitivas, donde los aprendizajes constituyan una de las variables principales de análisis, pero no la única.

Referencias

APRENDER 2016. *Informe de resultados*. En <http://www.bnm.me.gov.ar/giga1/documentos/EL005597.pdf>

- CTERA (2017). *Los operativos de evaluación "Aprender y Enseñar"*. Disponible en <http://mediateca.ctera.org.ar/files/original/6501cb036c7a188d2c505cd140397211.pdf>
- Documento Marco Aprender 2016. Ministerio de Educación y Deportes de la Nación. En http://educacion.gob.ar/data_storage/file/documents/las-claves-de-aprender-58bece1277804.pdf
- GONZÁLEZ OLGUÍN, E. (2015). *Tras una intervención global positiva, es momento de afrontar la diversidad en el sistema educativo*. Revista Educar en Córdoba N° 31, Junio 2015. En <http://revistaeducar.com.ar/2015/06/28/tras-una-intervencion-global-positiva-es-momento-de-afrontar-la-diversidad-en-el-sistema-educativo/>
- GUTIERREZ, G. (2011). *Dispositivos pedagógicos - Escuela y Enseñanza*. Notas para el análisis de una relación compleja. En S. La Rocca y G. Gutierrez (Comp.), *Escuela, Políticas y Formación Docente. Piezas en juego para una estrategia de transformación* (pp. 22-51). Córdoba: UEPC.
- GUTIERREZ, G., BELTRAMINO, L., CASTRO GONZÁLEZ, E., GONZÁLEZ OLGUÍN, E. & TOSOLINI, M. (2014). *El nivel secundario en Córdoba. Análisis de sus tendencias y transformaciones: 2003-2013*. Unión de Educadores de la Provincia de Córdoba. Córdoba: Alaya Editorial. <http://www.uepc.org.ar/conectate/publicacionespdf/instituto-UEPC-secundaria.pdf>
- MUÑOZ, V. (2013). *El derecho a la educación: una mirada comparativa. Argentina, Chile, Uruguay y Finlandia*. Oficina Regional de Educación para América Latina y el Caribe (OREALC/UNESCO Santiago). Disponible en <http://www.unesco.org/fileadmin/MULTIMEDIA/FIELD/Santiago/images/Estudio-comparativo-UNESCO-vernor-munoz-espanol.pdf>
- OCDE/CEPAL/CIAT/BID (2015). *Estadísticas tributarias en América Latina y el Caribe 1990-2013*. OECD Publishing, Paris. Disponible en http://dx.doi.org/10.1787/rev_lat-2015-3-en-fr
- SITEAL. Sistema de Tendencia Educativas de América Latina. Disponible en <http://www.siteal.iipe-oei.org/>
- TERCE. Tercer Estudio Regional Comparativo y Explicativo (TERCE) del Laboratorio Latinoamericano de Evaluación de la Calidad de la Educación (LLECE), OREALC/UNESCO Santiago. En <http://www.unesco.org/new/es/santiago/education/education-assessment-llece/third-regional-comparative-and-explanatory-study-terce/>

GONZALO GUTIERREZ

Facultad de Ciencias de la Comunicación – Cátedra Didáctica General,
Facultad de Filosofía y Humanidades – Cátedras de Didáctica General e Historia de la Educación Argentina,
Universidad Nacional de Córdoba (UNC).
Instituto de Capacitación e Investigación de los Educadores de la Provincia de Córdoba (ICIEC-UEPC).
(✉) gutierrezg61@yahoo.com.ar

Recibido: 13 de octubre de 2017.

Aceptado: 27 de octubre de 2017.

Publicado en línea: 1 de diciembre de 2017.
