

No por mucho pesarla, la vaca engorda. Reflexiones sobre evaluación educativa en un país ganadero como Uruguay

Não pesando muito, a vaca engorda. Reflexões sobre a avaliação educacional em um país pecuarista como o Uruguai

Not for constantly weigh it, the cow gets fat. Reflections on educational evaluation in a cattle-raising country like Uruguay

Gabriela Salsamendi¹

Andrés Peri²

Citação: SALSAMENDI, G. PERI, A. No por mucho pesarla, la vaca engorda. Reflexiones sobre evaluación educativa en un país ganadero como Uruguay. *Jornal de Políticas Educacionais*. V. 17, e92180. Dezembro de 2023.



<http://10.5380/jpe.v17i0.92180>

Resumen: En este artículo se presentan los principales hitos que han sido constituyentes del sistema de evaluación actual en Uruguay, en diálogo con una visión de la evaluación que evidencia el desacople entre la cantidad de mediciones y la mejora de los aprendizajes. En el primer apartado se relata la evolución del sistema de evaluación, su pasaje del papel a la aplicación en línea y su tránsito hacia una evaluación para el aprendizaje y no del aprendizaje. A continuación, se discute sobre esta tensión entre modelos de evaluación y se explica por qué Uruguay se ha orientado hacia las evaluaciones formativas en línea. En el segundo apartado, se indican las características principales de las evaluaciones disponibles para el cuerpo docente de educación primaria en la actualidad, sus ventajas y limitaciones. Luego se exponen una serie de argumentos en relación a la cantidad de evaluaciones que se realizan anualmente y cómo eso no ha repercutido en la mejora de los resultados de aprendizaje. Por último, se indican desafíos hacia adelante y se presenta una línea de trabajo con foco en evaluaciones adaptativas de aplicación autónoma que

¹ Magister en Educación. Docente del Departamento de Evaluación de Aprendizajes de la División de Investigación, Evaluación y Estadística de CODICEN. Asistente Académica de la Maestría en Educación, Sociedad y Política en FLACSO Uruguay. Montevideo. Uruguay. Orcid: <https://orcid.org/0009-0002-3756-0047> E-mail: a.g.salsamendi@gmail.com

² Doctor en Sociología. Director de la División de Investigación, Evaluación y Estadística de la Dirección Sectorial de Planificación Educativa de CODICEN. Montevideo. Uruguay. Orcid: <https://orcid.org/0009-0002-1941-1785> E-mail: aperi@anep.edu.uy

devuelven información a los centros educativos para la reflexión, el análisis y la planificación de los colectivos docentes. Allí se señalan los desafíos - y también las ventajas y los riesgos - de contar con este tipo de herramientas.

Palabras clave: Evaluación; Aprendizajes; Formativa.

Resumo: Este artigo apresenta os principais marcos constitutivos do sistema de avaliação atual no Uruguai, em diálogo com uma visão de avaliação que mostra a dissociação entre o número de medições e a melhoria da aprendizagem. A primeira seção descreve a evolução do sistema de avaliação, a sua passagem do papel para a aplicação online e a sua transição para uma avaliação de aprendizagem e não de aprendizagem. Em seguida, discute-se essa tensão entre os modelos de avaliação e explica-se por que o Uruguai se voltou para as avaliações formativas online. Na segunda seção, são indicadas as principais características das avaliações atualmente disponíveis para os professores do ensino básico, as suas vantagens e limitações. Em seguida, é exposta uma série de argumentos em relação ao número de avaliações que são realizadas anualmente e como isso não tem impactado na melhoria dos resultados de aprendizagem. Por fim, são apontados desafios futuros e apresentada uma linha de trabalho com foco em avaliações adaptativas de aplicação autônoma que devolvam informações aos centros educacionais para reflexão, análise e planejamento das turmas de ensino. Lá, são apontados os desafios - e também as vantagens e riscos - de se ter esse tipo de ferramenta.

Palavras-chave: Avaliação; Aprendizado; Formativa.

Abstract: This article presents the main milestones that have been constituents of the current evaluation system in Uruguay, in dialogue with a vision of evaluation that shows the decoupling between the number of measurements and the improvement of learning. The first section describes the evolution of the evaluation system, its passage from paper to online application and its transition towards an evaluation for learning and not of learning. Next, this tension between assessment models is discussed and it is explained why Uruguay has turned towards online formative assessments. In the second section, the main characteristics of the evaluations currently available to primary education teachers, their advantages and limitations are indicated. Then a series of arguments are exposed in relation to the number of evaluations that are carried out annually and how this has not had an impact on the improvement of learning results. Finally, challenges ahead are indicated and a line of work is presented with a focus on adaptive evaluations of autonomous applications that return information to educational centers for reflection, analysis and planning of the teaching groups. There, the challenges - and also the advantages and risks - of having this type of tool are pointed out.

Keywords: Assessment; Learning; Formative.

Introducción

Uruguay es un país ganadero en el que existe trazabilidad del ganado, es decir que es posible identificar el producto cárnico desde el nacimiento del animal hasta el final de comercialización de sus distintos subproductos. También existe la posibilidad de seguir la trayectoria de los estudiantes a partir de diferentes evaluaciones aplicadas durante la escolarización. Esta analogía resulta interesante para el campo educativo dado la necesidad de proteger las trayectorias individuales de cada estudiante en el pasaje por la escolaridad completa (incluso se podría incorporar información anterior o posterior a la escuela). A pesar de ello, este artículo pretende demostrar cómo los hitos que marcan el derrotero de estas trayectorias, particularmente las evaluaciones, no mueven la aguja de la balanza de los aprendizajes. Precisamente, por mucho que se pesen las vacas no engordan, por mucho que se midan los aprendizajes no mejoran.

Este artículo requirió una revisión documental de artículos e informes nacionales y de la región. En función de ella, se estructura en tres apartados. En el primero, se relata el desarrollo que ha tenido el campo de evaluación nacional y, particularmente, las transiciones fundamentales que otorgan los atributos característicos de las evaluaciones actuales: del papel a la aplicación en línea y de un modelo de evaluaciones **del** aprendizaje a otro de evaluaciones **para** el aprendizaje. En el segundo, se presentan las características más importantes de las evaluaciones que actualmente conforman el paisaje de evaluaciones para la educación primaria y se muestran los datos sobre cantidades de pruebas realizadas anualmente. La última sección de este apartado pone en tensión el tema de las cantidades de evaluaciones y la inamovilidad de los aprendizajes. Se pone de manifiesto esta tensión y se presentan posibles explicaciones. Finalmente, se esbozan los desafíos que como sistema de evaluación tenemos desde una mirada prospectiva. En especial, se expone acerca de una línea de evaluaciones incipiente, las evaluaciones adaptativas, que se enfocan en devolver información sustantiva a los colectivos de docentes de cada centro para que ellos analicen y planifiquen estrategias de mejora.

Transiciones. De evaluaciones en papel a evaluaciones en línea, de evaluaciones *del* aprendizaje a evaluaciones *para* el aprendizaje

Uruguay tiene una larga tradición en el terreno de la evaluación educativa en todos los tramos de la escolaridad obligatoria. En el caso particular de la educación primaria las evaluaciones estandarizadas tienen su primer antecedente en el año 1989. Estas primeras evaluaciones fueron desarrolladas desde la Oficina de Montevideo de la Comisión Económica para América Latina y el Caribe (CEPAL) que en esos años definió y llevó adelante la aplicación de pruebas estandarizadas para los estudiantes del último año de primaria, del último del ciclo básico de Media y en el último del ciclo superior de Media, es decir el sexto, el noveno y el doceavo año (MANCEBO, 2009, p. 8)

Estas primeras evaluaciones dieron origen a la creación, en el marco de la Administración Nacional de Educación Pública (ANEP), de unidades que específicamente atendieran el incipiente campo de la evaluación educativa nacional. Es así que en el marco del programa de Mejoramiento de la Calidad en la Educación Primaria (MECAEP) se conformó la Unidad de Medición de Resultados Educativos (UMRE) cuyo cometido era la evaluación de la educación primaria.

Luego de estas dos primeras fases - CEPAL y UMRE - la institucionalidad vuelve a cambiar, aunque se mantuvo dentro del propio sistema.

En una tercera fase se mantuvo el carácter interno de las evaluaciones en la ANEP pero éstas pasaron a ser desarrolladas desde el CODICEN de la ANEP, en lo que hoy es la División de Investigación, Evaluación y Estadística del CODICEN de la ANEP. Además el país se embarcó en evaluaciones internacionales como PISA de la OCDE (2003 y 2006) y SERCE de UNESCO (2006). (MANCEBO, 2009, p. 8).

La División de Investigación, Evaluación y Estadística (DIEE) fue la encargada de la realización de las evaluaciones nacionales de aprendizaje que se desarrollaron entre los años 1996 y 2013. En ese lapso de tiempo se aplicaron pruebas en papel y de aplicación controlada en educación inicial, primaria y media con la finalidad de informar al sistema acerca de los aprendizajes de los estudiantes. Las evaluaciones del último grado de primaria - el sexto - se realizaron en 1996, 1999, 2002, 2005, 2009 y 2013 en lectura, matemática y ciencias naturales. A esta serie se le sumaron las pruebas para el tercer grado de primaria (1998), para el tercer grado de media (1999), para inicial, 1ero. y 2do. (2001) y para sexto de media (2003).

Como ya se señaló estas pruebas tenían como objetivo la rendición de cuentas para informar a la política acerca de los avances, o no, de los aprendizajes. En 2013, es la última oportunidad en que la DIEE se encarga de esta prueba y es la última prueba que se compara con las series anteriores dado que con la creación del Instituto Nacional de Evaluación Educativa (INEEd) las pruebas estandarizadas pasan a su órbita y se discontinúa la serie.

Por otro lado, en paralelo a las evaluaciones nacionales de aprendizaje, se venían desarrollando otras evaluaciones complementarias a las primeras. Desde la DIEE se estaba trabajando en un nuevo tipo de evaluaciones: las evaluaciones formativas en línea.

De hecho, con la creación del Plan Ceibal, en 2007, la DIEE se propone avanzar hacia la primera de las transiciones que actualmente otorgan uno de los atributos de las evaluaciones de la educación primaria en Uruguay: pasar de evaluaciones en formato papel a evaluaciones en línea. Es preciso señalar que para que esto sucediera fue indispensable el reparto de una computadora portátil a cada estudiante y a cada docente de primaria, además de proveer de conectividad a internet a todos los centros educativos.

Con estas condiciones dadas se desarrolla una plataforma de evaluación que permite la aplicación de evaluaciones en línea de manera autónoma en el segundo grado de la educación primaria. Estas evaluaciones fueron desarrolladas con actividades de múltiple opción que evaluaban las áreas de lectura, matemática y ciencias naturales. La decisión de enfocarse en segundo año respondía a que en 2007 se había aplicado una prueba en el primer grado y no todos los niños eran capaces de leer autónomamente y contestar las actividades.

Al año siguiente, se volvió a ofrecer la evaluación en línea para segundo año en las mismas áreas de conocimiento, pero la apuesta se redobló al liberar actividades de la Evaluación Nacional de Aprendizajes de Sexto de 2009, como prueba en línea para ese grado escolar. En 2010 la participación de los estudiantes de segundo año y de sexto año - 70 y 75 % de cobertura respectivamente - evidenció la apropiación paulatina de una nueva herramienta que venía a cambiar el paisaje en las aulas uruguayas: la evaluación en línea (LUACES, 2014, p. 117).

En 2011, se vuelve a revisar la decisión de los grados de aplicación y se observa que en segundo grado muchos niños todavía no habían adquirido la lectura autónoma por lo que no se podía inferir si una contestación errada frente a una actividad de cualquiera de las tres áreas, se debía a que no sabían resolverla o a que no sabían leerla. Por lo que, a partir de ese año, la evaluación en línea se disponibilizó para los grados de tercero a sexto de educación primaria en las mismas áreas que antes. Ese año se optó por elaborar una prueba destinada a los estudiantes de tercero y cuarto grado, y otra prueba para quinto y sexto. Por otra parte, en ese ciclo se incorporó la idea de que las pruebas iban a estar disponibles en los portales de los docentes a mitad del año lectivo, de modo de que hubiese tiempo para trabajar a partir de la información que generaba la prueba en la segunda mitad del año.

Esta es la segunda transición sustantiva en términos de pruebas: pasar de la aplicación de evaluaciones estandarizadas enfocadas en reportar resultados para la toma de decisiones de política educativa al desarrollo de evaluaciones de aplicación autónoma que devolvieran información a los centros. Esta información era habilitada para generar análisis en el colectivo de docentes de una institución, por lo que se dispuso el pago de una sala de docentes fuera de horario a estos efectos. Esto no significa, a nivel de país, resignar las primeras para aplicar exclusivamente las segundas. Por el contrario, la visión de complementariedad entre los objetivos de ambos tipos de evaluación es la que más

sentido cobró al momento de pensar las evaluaciones como palancas para movilizar aprendizajes.

Ya en 2012, se normaliza la aplicación de evaluaciones en línea formativas dirigidas a los cuatro grados superiores de la educación primaria con una prueba para cada grado en cada área de conocimiento evaluada. En total se elaboraron doce pruebas con actividades propias para el grado, con actividades comunes a grados consecutivos y transversales a todos los grados evaluados. Estas últimas, favorecen una visualización con foco en los avances entre estudiantes de distintos grados. También en esta edición se incorporan las actividades de respuesta abierta, es decir, en la que los estudiantes elaboran una respuesta que, posteriormente, es codificada por el docente.

Desde 2013 y hasta 2020 se mantuvo el formato anteriormente descrito. Sin embargo, en 2020, debido a la situación de pandemia, se decidió pasar de pruebas que tenían entre 18 y 24 actividades por área a pruebas más cortas. Las pruebas desde 2020 hasta 2023 dirigidas a los cuatro grados superiores de primaria están conformadas por cuatro módulos de diez actividades cada uno. Los módulos están estructurados por el área de conocimiento que evalúan. Hay un módulo de lectura, uno de matemática, uno de ciencias naturales y otro multiárea que presenta actividades de las tres áreas de conocimiento articuladas en torno a un tema común.

Lo más interesante de esta última transición es la instalación de una nueva visión frente a las evaluaciones. La evaluación en línea es una evaluación en clave formativa. No permite decir cuánto saben los estudiantes de un grado escolar en un área de conocimiento, sin embargo, habilita intervenciones docentes que atienden a las respuestas dadas por sus estudiantes enfrentados a un conjunto de actividades desafiantes. En ese sentido, se trata de una evaluación que entra a las aulas a complementar otras que elabora y aplica cada docente. Su virtud con respecto a las que crea el docente de manera individual es la de propiciar la conversación entre docentes sobre el mismo referente conceptual (ALONSO; PERI, 2013, p. 280-301).

Esta visión de las evaluaciones está muy influenciada por ideas de Michael Fullan y de María Langworthy (2014, p. 56-57) en relación a considerarla como un elemento que fomenta culturas colaborativas en tanto es un instrumento de diálogo entre docentes - rompiendo el aislamiento del aula - y un recurso que apoya la enseñanza a mitad de un año lectivo.

Características de las evaluaciones disponibles para los docentes en la actualidad

Las pruebas que en la actualidad están disponibles para la educación primaria en Uruguay, son ofrecidas por dos instituciones. Por un lado, están las que desarrolla el INEEd que libera una prueba conformada por un conjunto de actividades liberadas de Aristas, para aplicación autónoma y en línea, con el nombre Aristas en clase. Estas pruebas permiten evaluar lectura y matemática en el tercer grado y en el sexto grado de primaria para comparar los resultados de un grupo en particular con los resultados nacionales en la aplicación controlada. Estas pruebas son voluntarias, sin embargo, no han alcanzado grandes cantidades de usuarios.

Por otro lado, desde la DIEE anualmente se desarrollan y habilitan las evaluaciones formativas en línea descritas en el apartado anterior. Actualmente estas pruebas están conformadas por cuatro módulos de diez actividades cada uno que se aplican de tercero a sexto en lectura, matemática y ciencias naturales. La plataforma de aplicación ha incorporado nuevos formatos de ítems - además de los tradicionales de múltiple opción y abiertos de desarrollo - que permiten utilizar menús desplegables, completar espacios en blanco, arrastrar y soltar, entre otros.

Las pruebas se publican en los portales de docentes, directores e inspectores y son acompañadas de documentos de aportes que presentan marcos teóricos y decisiones metodológicas que permiten analizar los ítems. Vale señalar que al tratarse de módulos con pocas actividades ha sido necesario hacer foco en algunos contenidos o competencias y no tienen amplia cobertura curricular. Estos focos se definen para cada ciclo anual en coordinación entre los equipos de evaluadores, supervisores y formadores. Por otro lado, al tratarse de una prueba con solo diez actividades por módulo se estimula la retroalimentación inmediata, es decir, el docente conoce las actividades antes de aplicarlas, propone la aplicación e inmediatamente luego de finalizarla, analiza los resultados de algunas actividades seleccionadas con sus estudiantes a fin de reflexionar sobre las respuestas dadas. De acuerdo a Sadler (2010, p. 2-13) la retroalimentación ofrece la posibilidad de que los estudiantes tomen conciencia de la brecha que existe entre el nivel en el que se encuentran y el que se espera que alcancen.

En el último año completo, 2022, se aplicaron 514.322 pruebas del ciclo de evaluaciones en línea formativas entre tercero y sexto grado. Esto evidencia el grado de

apropiación de los docentes frente a este instrumento. Incluso el año anterior, en contexto de pandemia, también fueron aplicadas cerca de medio millón de pruebas.

A continuación, se presentan los datos acerca de la cantidad de pruebas aplicadas por grado en cada ciclo de evaluaciones formativas en línea y las características fundamentales de cada ciclo.

Cuadro 1 –Evolución de pruebas formativas en línea para Primaria

Año de aplicación	Pruebas por grados evaluados	Áreas evaluadas por módulo	Cantidad de ítems por prueba o módulo	Tipos de actividades	Cantidad de aplicaciones
2009	Una prueba para 2do	Lectura, matemática y ciencias naturales	Entre 18 y 24 actividades	Múltiple opción	87.000
2010	Una prueba para 2do y otra para 6to	Lectura, matemática y ciencias naturales	Entre 18 y 24 actividades	Múltiple opción	270.000
2011	Una prueba para 3ero y 4to y otra para 5to y 6to	Lectura, matemática y ciencias naturales	Entre 18 y 24 actividades	Múltiple opción	457.000
2012	Una prueba por grado: 3ero, 4to, 5to y 6to.	Lectura, matemática y ciencias naturales	Entre 18 y 24 actividades	Múltiple opción y abiertas de desarrollo	480.000
2013	Una prueba por grado: 3ero, 4to, 5to y 6to	Lectura, matemática y ciencias naturales	Entre 18 y 24 actividades	Múltiple opción y abiertas de desarrollo	793.000 ³
2014	Una prueba por grado: 3ero, 4to, 5to y 6to	Lectura, matemática y ciencias naturales	Entre 18 y 24 actividades	Múltiple opción y abiertas de desarrollo	852.000 ⁴
2015	Una prueba por grado: 3ero, 4to, 5to y 6to	Lectura, matemática y ciencias naturales	Entre 18 y 24 actividades	Múltiple opción y abiertas de desarrollo	542.331
2016	Una prueba por grado: 3ero, 4to, 5to y 6to	Lectura, matemática y ciencias naturales	Entre 18 y 24 actividades	Múltiple opción y abiertas de desarrollo	603.235
2017	Una prueba por grado: 3ero, 4to, 5to y 6to	Lectura, matemática y ciencias naturales	Entre 18 y 24 actividades	Múltiple opción y abiertas de desarrollo	539.084
2018	Una prueba por grado: 3ero, 4to, 5to y 6to	Lectura, matemática y ciencias naturales	Entre 18 y 24 actividades	Múltiple opción y abiertas de desarrollo	537.761

³ Esta cifra considera aplicaciones finalizadas e iniciadas.

⁴ Esta cifra considera aplicaciones finalizadas e iniciadas.

2019	Una prueba por grado: 3ero, 4to, 5to y 6to	Lectura, matemática y ciencias naturales	Entre 18 y 24 actividades	Múltiple opción, abiertas de desarrollo, correlación, menú desplegable, cerrado con justificación, espacios en blanco.	542.923
2020	Una prueba por grado: 3ero, 4to, 5to y 6to	Lectura, matemática, ciencias naturales y multiárea	10 actividades	Múltiple opción, abiertas de desarrollo, correlación, menú desplegable, cerrado con justificación, espacios en blanco.	483.879
2021	Una prueba por grado: 3ero, 4to, 5to y 6to	Lectura, matemática, ciencias naturales y multiárea	10 actividades	Múltiple opción, abiertas de desarrollo, correlación, menú desplegable, cerrado con justificación, espacios en blanco.	455.748
2022	Una prueba por grado: 3ero, 4to, 5to y 6to	Lectura, matemática, ciencias naturales y multiárea	10 actividades	Múltiple opción, abiertas de desarrollo, correlación, menú desplegable, cerrado con justificación, espacios en blanco.	514.322
2023	Una prueba por grado: 3ero, 4to, 5to y 6to	Lectura, matemática, ciencias naturales y multiárea	10 actividades	Múltiple opción, abiertas de desarrollo, correlación, menú desplegable, cerrado con justificación, espacios en blanco.	En ejecución

Fuente: elaboración propia.

Ahora bien, en un país en donde se aplican pruebas estandarizadas de rendición de cuentas nacionales, regionales e internacionales, pero además se desarrollan y aplican pruebas formativas en línea desde hace más de una década - en las que se alcanzan cifras altísimas de cobertura-, un país en donde se ha transitado hacia la coexistencia y la complementariedad de distintos modelos de evaluaciones - **del** aprendizaje y **para** el aprendizaje -, ¿han mejorado los aprendizajes? La evidencia tanto nacional, como la que

se observa en evaluaciones regionales e internacionales en las que participa Uruguay, muestra que los aprendizajes no han variado significativamente, tal como veremos en el siguiente apartado. Mover los aprendizajes de los estudiantes es muy difícil; no por mucho pesarla, la vaca engorda.

Evaluaciones como impulsores de los aprendizajes

El INEEd, por mandato legal, produce regularmente un informe sobre el estado de la educación en Uruguay. Precisamente, en el Informe sobre el estado de la educación en Uruguay 2015-2016 se presentaron datos comparados de las últimas evaluaciones nacionales de aprendizajes y un gráfico que indica el promedio y la distribución del puntaje de cada área evaluada según las distintas evaluaciones entre 1996 y 2013 (INEED, 2017, p. 102).

Gráfico 1 – Promedio y distribución del puntaje en cada área evaluada según año



Fuente: INEEd (2017).

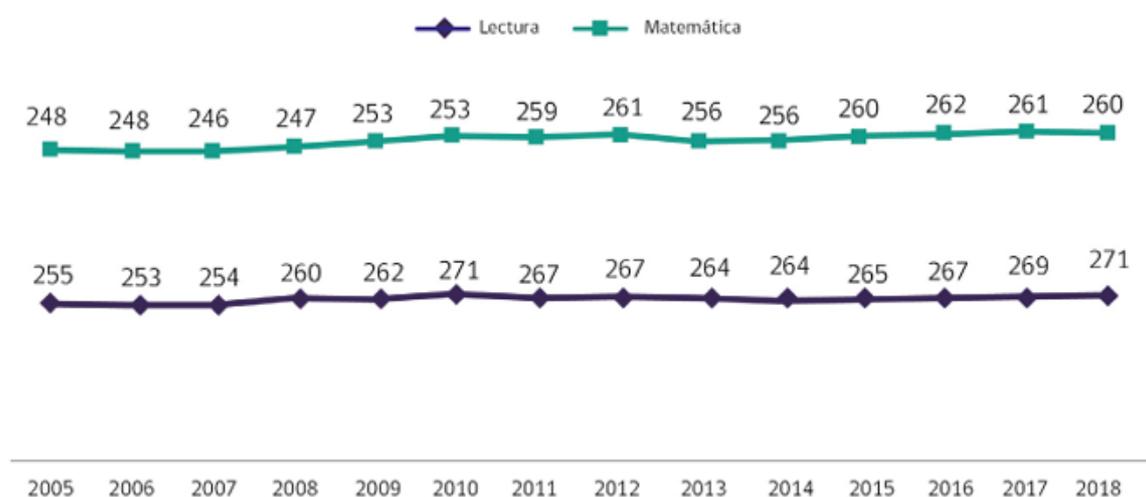
En el mismo informe se señala que

[...] las diferencias en los puntajes promedio entre 1996 y 2009 no son significativas en lectura; en matemática se observa una mejora en 2002 respecto de 1996, para luego registrarse una caída en 2009; en ambas áreas los resultados de 2013 son notoriamente inferiores al resto de los años; y en ciencias las diferencias entre 2009 y 2013 no son relevantes. En términos generales, ello parece indicar que no ha habido mejora en los resultados. (INEEd, 2017, p. 101).

De hecho, esta misma situación ocurre si se analiza la evolución del puntaje de Uruguay en PISA. Desde 2003 a 2018 Uruguay se ha desempeñado siempre en una franja de puntajes entre 409 y 437 puntos, en las tres áreas: lectura, matemática y ciencias naturales. Esta situación no es exclusiva de Uruguay, sino que para el conjunto de países que participaron de la evaluación PISA, entre 2000 y 2018, la tasa de crecimiento de aprendizaje es cero (RIVAS; SCASSO, 2019, p. 117-133; WILLMS, 2018, p. 9-12)

Incluso si se analizan resultados de otras evaluaciones nacionales de aprendizaje en otros países de la región también se puede observar que la aguja de los resultados de aprendizajes prácticamente no se mueve. Véase a continuación el gráfico que presenta el promedio nacional en lectura y en matemática en SIMCE de 4to básico entre 2005 y 2018:

Gráfico 2 – Resultados Simce 4º básico 2005-2018



Fuente: Agencia de Calidad de la Educación, Chile (2019).

Todos estos elementos permiten mostrar que es muy difícil mover la aguja de los aprendizajes y a continuación, se esbozan algunas posibles explicaciones frente a este estancamiento.

Lo primero sobre lo que es preciso reflexionar es cómo medimos en los países de la región los resultados de aprendizajes o, dicho de otra manera, qué medimos cuando medimos.

En general, las mediciones de aprendizajes se hacen con base en información que arroja la aplicación de pruebas estandarizadas - mayoritariamente enfocadas en medir lectura y matemática - de aplicación controlada sobre muestras representativas de la población. Sin embargo, esas pruebas son el producto de otras determinaciones: qué ítems las conforman, quiénes y sobre qué base se elaboran, cómo se arman las pruebas y los operativos para llevarlas adelante, cómo se conforman las muestras, entre otros elementos. Estos diseños, a su vez, pueden atender a distintas teorías al momento de los análisis: teoría clásica de los test (TCT), teoría de respuesta al ítem (TRI), o más recientemente a la aplicación de TRI en test adaptativos informatizados (TAI).

Todos estos elementos permiten generar una visión sobre los aprendizajes, visión que da información sobre respuestas a actividades alineadas a un constructo. La pregunta que sobrevuela es si los resultados de la educación son solo estos o existen otros productos de la educación que no hemos sido capaces de medir. Dicho de otra manera, la educación genera otros resultados, otros productos, otros saberes que no siempre son medibles e incluso no valdría la pena el esfuerzo de medirlos.

Lo que sí es sabido es que el desarrollo de pruebas estandarizadas y la aplicación regular de ellas, permite la comparación de distintas cohortes en el tiempo. Hasta este punto es lo que los países hemos podido hacer con lo que se ha desarrollado. Hay que considerar que cada evaluación termina siendo una operación de medición agregada, se mide en un grado lo que cada estudiante ha incorporado desde el inicio de su escolaridad y también desde sus saberes y experiencias extraescolares. Incluso sabemos que inciden en esta medición otros aspectos: sanitarios, culturales, de motivación, de gestación, entre otros.

Lo que hacemos cuando evaluamos en pruebas nacionales de aprendizajes, es captar la reacción de los estudiantes frente a un reactivo o actividad. Ese conjunto de actividades intenta captar una habilidad subyacente pero no todas las habilidades ni todos los saberes que se generan en la escuela (ni en la vida). Los propios formatos de las pruebas - en papel o en línea - excluyen la posibilidad de proponer actividades que releven otras habilidades que pueden evaluarse lateralmente pero no de manera directa: observación científica, curiosidad, creatividad, programación, uso de instrumentos, etc.

Estos también son conocimientos que se construyen en la escuela. Cada vez más los saberes escolares se han expandido en múltiples direcciones y, aunque creemos que pueden colaborar en la mejora de las áreas tradicionalmente evaluadas, no logramos captarlos con las evaluaciones estandarizadas. Los docentes sí pueden captarlos en la dinámica de la cotidianeidad del aula.

Por otra parte, todas las actividades de evaluación requieren de otra habilidad que es independiente del área evaluada que es la lectura. Si el estudiante logra comprender el enunciado, la consigna, las tablas, gráficos, dibujos que pueda tener la actividad, podrá responder la propuesta, de lo contrario no va a tener esa posibilidad incluso si conoce la respuesta de lo que se le está solicitando.

Sin embargo, es posible pensar en movilizar aprendizajes en cada escuela para mover la aguja localmente antes de mover la *gran* aguja. Es desde las escuelas que es posible generar los cambios atendiendo a la realidad contextual del centro, al conocimiento profesional de los colectivos docentes, a los aportes de la comunidad y a los recursos disponibles.

Las evaluaciones de tercera generación, TAI, permiten captar los niveles individuales de conocimiento de cada estudiante. Esto permitiría, en primer lugar, devolver información sobre el nivel de desempeño de cada estudiante al docente y al director para generar estrategias de enseñanza en clave de personalización.

En segundo lugar, esta información podría impulsar políticas de focalización que superen la visión de compensar al centro educativo y avancen en la dirección de compensar al estudiante que lo requiera.

En tercer lugar, las evaluaciones adaptativas proponen una mirada distinta acerca de la estructuración de los formatos escolares de aula monogrado ya que ofrecen evidencia sobre los niveles de desempeño que desacomodan esa estructura: hay estudiantes en casi todos los niveles de desempeño en cada grupo y hay estudiantes con el mismo nivel de desempeño en distintos grados escolares.

En Uruguay, la evaluación adaptativa (TAI) que se desarrolla desde la DIEE lleva el nombre de SEA+. Este instrumento se ha venido gestando desde 2011 cuando la tecnología que aportaba el Plan Ceibal permitió soñar con este salto cualitativo en términos de evaluaciones. En 2012 se aplicó el primer piloto de actividades para calibrar los ítems en una misma métrica, sin embargo, la falta de marcos curriculares que indicaran progresión en los constructos a evaluar llevó al proyecto a una pausa.

Avanzada la década de 2010, se elaboraron perfiles de egreso y progresiones curriculares que volvieron a poner en agenda el proyecto SEA+. Es así que en 2018 se realizó una prueba de concepto en 44 escuelas uruguayas con el objetivo de conocer si para los docentes, directores e inspectores este proyecto tenía sentido pedagógico en las escuelas. En esa etapa se acompañó de cerca el proyecto recorriendo los territorios. Con sorpresa se observaron cambios en las prácticas tanto a nivel de aula, como de escuela.

A nivel de aula, se pudo advertir que los docentes tomaban en cuenta los niveles de desempeño que les devolvía SEA+ en dos sentidos: primero, al momento de proponer una consigna de trabajo individual se buscaba que cada estudiante pudiese trabajar a su nivel, pero con la mirada puesta en acompañarlo hacia el nivel siguiente. Esto supuso una planificación en niveles con consignas adecuadas y con el propósito de avanzar gradualmente. Por otro lado, también se observaron propuestas por equipos en donde en algunas ocasiones los equipos eran conformados por estudiantes según su nivel y en otros se buscaba equilibrar los niveles de cada equipo. Todas estas decisiones de cada docente atendían a distintos objetivos en relación con las actividades planteadas.

A nivel de centro educativo, se pudo observar que los docentes analizaban los datos que ofrecía el instrumento al final del año lectivo pero que lo más relevante sucedía al comenzar el siguiente año. Hay que recordar que la información que recoge SEA+ al final de cada año lectivo, migra en la plataforma de los docentes para que, al comienzo del siguiente año, cuenten con los niveles de desempeños alcanzados por sus estudiantes en el año anterior. En ese sentido, SEA+ movilizó a los colectivos para llevar adelante nuevas prácticas escolares. Fuimos testigos de experiencias de desestructuración de grupos o de grados escolares y de proyectos de ciclo en los que los estudiantes, independientemente de su grado de origen, se reagrupaban con otros estudiantes de otros grados con los que compartían el mismo nivel de desempeño.

En 2019, se decidió ampliar la cantidad de centros que podían participar y se invitó a todas las escuelas de uno de los departamentos del país. Ese año se anotaron para aplicar SEA+, 159 escuelas. Este dato muestra el interés por ser parte de este proyecto dado que les ofrecía a los colectivos docentes una información que resultaba relevante a los efectos educativos. Al final de ese año se resolvió que para el próximo se iba a comenzar la primera etapa de universalización invitando a las escuelas de la mitad del país. Sin embargo, 2020 nos sorprendió con un escenario de pandemia en donde el avance

hacia la mitad del país fue complejo por lo que en 2021 y 2022 se volvió a trabajar con las escuelas de ese medio país.

Cuadro 2. Implementación de fase inicial del proyecto 2017 - 2022

Año	Escuelas inscriptas	Departamentos
2018	44	Canelones Este, Canelones Centro, Canelones Oeste.
2019	159	Canelones Este, Canelones Centro, Canelones Oeste.
2020	423	Montevideo Oeste, Artigas, Canelones Este, Canelones Centro, Canelones Oeste, Colonia, Flores y Salto.
2021	544	Montevideo Oeste, Artigas, Canelones Este, Canelones Centro, Canelones Oeste, Colonia, Flores y Salto. Además participaron escuelas del Proyecto MATEC.
2022	473	Montevideo Oeste, Artigas, Canelones Este, Canelones Centro, Canelones Oeste, Colonia, Flores y Salto. Además participaron escuelas del Proyecto MATEC.

Fuente: Elaboración propia.

SEA+ nos ha mostrado una nueva visión sobre las evaluaciones. Nos demuestra cómo la información puede ser utilizada para efectuar cambios en las prácticas y para movilizar a los colectivos docentes en torno a un rumbo. Si bien no por mucho medir, ni por mucho pesar a la vaca, los aprendizajes van a mejorar, este tipo de evaluaciones con foco en los centros representan una oportunidad de cambio en el campo de las evaluaciones.

Las evaluaciones como impulsores de la mejora de aprendizaje, suponen un cambio de visión sobre ellas. Este cambio no implica dejar de medir, ni dejar de aplicar evaluaciones nacionales con foco en informar para el desarrollo o el monitoreo de políticas educativas. En cambio, implican pensar en la complementariedad de estas evaluaciones con otras en donde los protagonistas sean las comunidades educativas.

Si bien estas evaluaciones adaptativas no son la panacea, ni tienen el cometido de mejorar los aprendizajes de los estudiantes, es posible ver en ellas un camino nuevo ligado a la autonomía de la gestión escolar y a la profesionalización de los docentes. Estas evaluaciones, como ya se señaló, están entregando información directamente a los centros educativos: docentes, directores e inspectores. Información que contribuye a reflexionar sobre las prácticas educativas y sobre lo que es posible hacer desde la escuela

para andamiar los procesos de aprendizaje personales que transita cada estudiante. Interpela la idea de escuela graduada o la concepción de que todos los estudiantes que comparten el aula aprenden lo mismo, al mismo tiempo y del mismo modo y abre la oportunidad de repensar la estructura escolar a la luz de la diversificación de las prácticas de enseñanza.

Conclusiones

Este artículo ha recorrido los principales hitos constituyentes del sistema de evaluación de aprendizajes en Uruguay. Se ha estructurado en tres apartados.

En el primer apartado se han presentado las dos transiciones que otorgan la identidad al sistema de evaluación de aprendizajes uruguayo: de la aplicación en papel a la aplicación en línea y de evaluación **del** aprendizaje a evaluaciones **para** el aprendizaje. En el segundo, se señalaron las características de las evaluaciones que disponibiliza la ANEP para el magisterio uruguayo y se expusieron algunos argumentos que han mostrado cómo la cantidad de mediciones o evaluaciones que se hacen, no tiene como consecuencia la mejora de los aprendizajes. No por mucho pesarla, la vaca engorda. En el tercero y último, se presentó una nueva línea de evaluaciones: las evaluaciones adaptativas informatizadas que devuelven información a los docentes y directores de los centros educativos para generar instancias de reflexión y para promover estrategias locales de mejora basadas en la información como evidencia.

Si bien no es posible afirmar que este instrumento -SEA+ - va a mover la aguja de los aprendizajes medidos de la misma manera en que medimos ahora, la experiencia nos ha demostrado que resulta movilizador de otros engranajes en los centros educativos. Estos instrumentos que informan a los colectivos, muchas veces resultan reiterativos para los docentes dado que reportan información que ya conocen de los estudiantes; la ventaja está en dos asuntos: primero, corroboran la visión de los docentes - aunque en algunos casos también se llevan sorpresas - y segundo, habilitan el diálogo entre docentes, promueven culturas colaborativas que instan a pensar nuevas formas de hacer escuela y ponen en el centro a cada estudiante de manera personalizada.

Finalmente, estos instrumentos adaptativos propician el uso de información en escala vertical, la información sobre los niveles de desempeño alcanzados por un alumno en particular en los distintos grados escolares conforma una suerte de historia escolar y

ofrece evidencia de qué pudo hacer ese estudiante al final de tercero, cuarto, quinto y sexto año, cuál fue su ganancia de aprendizaje interanual, dónde pudo haber estancamientos o mayores avances, entre otros tipos de información como evidencia. En los centros educativos se han generado experiencias innovadoras que suponen el rompimiento de los grupos y de los grados tradicionales para atender a los niveles del estudiantado, proyectos institucionales en clave de personalizar las prácticas, planificaciones diversificadas con consignas ajustadas a cada estudiante, espacios de trabajo individual, tutorías entre pares de alumnos, entre otras. Por lo que esperamos que estas evaluaciones – pesar la vaca – arrojen evidencia que provoquen prácticas de enseñanza personalizada que colaboren en la mejora de los aprendizajes – y que la vaca engorde-.

Referencias

AGENCIA DE CALIDAD DE LA EDUCACIÓN. **Evaluaciones Nacionales e Internacionales de Aprendizaje Período 2004-2018 Nueva edición con datos hasta 2018**. Santiago de Chile. 2019.

ALONSO, Nicolás; PERI, Andrés. **La evaluación de los aprendizajes en línea en clave de innovación**. ANEP: Uruguay. 2013.

ANEP, PISA. **Uruguay en la evaluación PISA**. Primeros resultados del ciclo 2018. Montevideo, 2019.

FULLAN, Michael; LANGWORTHY, María. **Una rica veta: cómo las nuevas pedagogías logran el aprendizaje en profundidad**. 2014.

INEEd. **Informe sobre el estado de la educación en Uruguay 2015-2016**, INEEEd, Montevideo. 2017.

LUACES, Oscar. **Evaluaciones en línea: una innovación educativa en el marco del programa de una computadora por alumno**. ANEP, 2014.

MANCEBO, María Ester. El Instituto Universitario de Educación y el Instituto Nacional de Evaluación Educativa: una ventana de oportunidad para el desarrollo educativo en Uruguay. **VIII Jornadas de Investigación “El futuro del país en debate”**, 2009, p. 1-16.

RIVAS, Axel; SCASSO, Martín Guillermo. La segunda dimensión de la calidad educativa: Un estudio comparado en siete países de América Latina. **Bordón: Revista de pedagogía**, 2019, vol. 71, no 4, p. 117-133.

SALSAMENDI, G. PERI, A. No por mucho pesarla, la vaca engorda. Reflexiones sobre evaluación educativa en un país ganadero como Uruguay

SADLER, D. Royce. Beyond feedback: Developing student capability in complex appraisal. **Assessment & evaluation in higher education**, 2010, vol. 35, no 5, p. 535-550, 2010.

WILLMS, Douglas. **Las brechas de aprendizaje: uso de datos para informar a la política** Documento Informativo. Unesco, No. 54 Septiembre de 2018.

Recebido em Agosto de 2023
Aprovado em Novembro de 2023
Publicado em Dezembro de 2023



Volume 17

e92180

11 de Dezembro de 2023



O Copyright é retido pelo/a autor/a (ou primeiro co-autor) que outorga o direito da primeira publicação ao **Jornal de Políticas Educacionais**. Mais informação da licença de Creative Commons encontram-se em <https://creativecommons.org/licenses/by-nc-nd/4.0>. Qualquer outro uso deve ser aprovado em conjunto pelo/s autor/es e pelo periódico.

JORNAL DE POLÍTICAS EDUCACIONAIS é uma publicação do Núcleo de Políticas Educacionais do Setor de Educação da Universidade Federal do Paraná – NuPE/UFPR, em consórcio com a Linha de Pesquisa em Políticas Educacionais do Programa de Pós-Graduação em Educação – PPGE/UFPR, que aceita colaboração, reservando-se o direito de publicar ou não o material espontaneamente enviado à redação. As colaborações devem ser enviadas ao NuPE/UFPR, conforme orientações contidas nas páginas do periódico na internet: <http://revistas.ufpr.br/jpe>.

INDEXAÇÃO:

BASE DE DADOS

Sumários.Org

Google Scholar

BASE

Dimensions

Miar

DIRETÓRIOS

Scielo Educ@

Diadorim

DOAJ

ÍNDICES

Index Copernicus

Cite Factor

PORTAIS

LiVre

Capes

Science Open

World Wide Science

Erih Plus

Latindex

EZB

ROAD

Journal 4-free

(Periódico integralmente disponível apenas em via eletrônica)

Jornal de Políticas Educacionais / Núcleo de Políticas Educacionais da Universidade Federal do Paraná – NuPE/UFPR – v.1, n. 1 (1º semestre de 2007) – Curitiba: NuPE/UFPR.

Volume 17, e92180 – Dezembro de 2023

ISSN 1981-1969

1. Educação – Periódicos. 2. Política Educacional – Periódicos. I. NuPE/UFPR

Comitê Editorial:

Adriana Aparecida Dragone Silveira (UFPR-Brasil)

Gabriela Schneider (UFPR-Brasil)

Conselho Editorial:

Ana Lorena de Oliveira Briel (UFPR - Brasil), Andréa Barbosa Gouveia (UFPR - Brasil), Ângela Maria Martins (FCC - Brasil), Ângelo Ricardo de Souza (UFPR - Brasil), Antonia Almeida Silva (UEFS - Brasil), Cassia Alessandra Domiciano (UFPR - Brasil), Cesar Tello (UNTREF - Argentina), Claudia Regina Baukat Silveira Moreira (UFPR - Brasil), Cristiane Machado (Unicamp - Brasil), Daniela de Oliveira Pires (UFPR - Brasil); Elisângela Scaff (UFPR - Brasil); Elton Luiz Nardi (UNOESC - Brasil), Fernanda Saforcada (UBA - Argentina), Isaac Paxé (ISCED - Angola), Gladys Beatriz Barreyro (USP - Brasil), Gilda Cardoso Araújo (UFES - Brasil), Gustavo Enrique Fischman (ASU - Estados Unidos), Janete Maria Lins de Azevedo (UFPE - Brasil), Jefferson Mainardes (UEPG - Brasil), João Ferreira de Oliveira (UFG - Brasil), Jorge Alarcon Leiva (UTalca - Chile), Jorge Manuel Gorostiaga (UNSAM - Argentina), Juca Gil (UFRGS - Brasil), Luciana Rosa Marques (UFPE - Brasil), Marcos Alexandre dos Santos Ferraz (UFPR - Brasil) Marcia Aparecida Jacomini (Unifesp - Brasil), Maria Dilnéia Espíndola Fernandes (UFMS - Brasil), Natalia Oliveira Woolley (UCLA - Estados Unidos), Ney Cristina Monteiro de Oliveira (UFPA - Brasil), Nicolás Bentancur, (UDELAR - Uruguai), Nora Krawczyk (Unicamp - Brasil), Pedro Flores-Crespo (UAQ - México) Rodrigo da Silva Pereira (UFBA - Brasil), Robert Verhine (UFBA - Brasil), Rosana Cruz (UFPI - Brasil), Rubens Barbosa Camargo (USP - Brasil), Sebastián Donoso Díaz (UTalca - Chile), Theresa Adrião (UNICAMP - Brasil), Vera Maria Vidal Peroni (UFRGS - Brasil).

Créditos e Agradecimentos:

Revisão de Língua Portuguesa, Abstract e Resumen: Programa de apoio às publicações científicas periódicas da UFPR

Revisão de normatização: Giovanna Liz Cabral de Oliveira

Arte e diagramação: Tiago Tavares (tiagotav@gmail.com)

Jornal de Políticas Educacionais
Universidade Federal do Paraná
Setor de Educação
Núcleo de Políticas Educacionais – NuPE/UFPR
Avenida Sete de Setembro, 2645
2º andar, Sala 213
80.230-010 – Curitiba – PR – Brasil
jpe@ufpr.br
<http://revistas.ufpr.br/jpe>