



Programa de Promoción de la Reforma Educativa en América Latina y el Caribe
Partnership for Educational Revitalization in the Americas

Para comprender las evaluaciones educativas Fichas didacticas

Ficha 12

**¿cómo usar las evaluaciones para mejorar
la educación?**

Pedro Ravela

¿cómo usar las evaluaciones para mejorar la educación?

los resultados como herramienta de aprendizaje y mejora

Como complemento de las fichas 10 y 11, la ficha 12 tiene como objetivo mostrar enfoques para el uso de los resultados de las evaluaciones estandarizadas alternativos a los rankings y consecuencias "fuertes".

En esta ficha se busca mostrar experiencias dirigidas a utilizar los resultados de manera tal que contribuyan a mejorar el trabajo de escuelas y docentes y los logros educativos de los estudiantes a través del aprendizaje y el desarrollo profesional.

Para ello se vuelve sobre el tema de los usos formativos de las evaluaciones y el papel

que la información resultante de ellas puede tener para cuatro audiencias principales -- educadores, familias, autoridades y opinión pública -- en la medida en que la información se entregue de manera adecuada, evitando las distorsiones y errores analizados en fichas anteriores.

A modo de introducción al tema de esta ficha, el recuadro 1 pretende destacar la importancia de los usos formativos de las evaluaciones, desde la visión del destacado especialista chileno Ernesto Schiefelbein.

Recuadro 1

Uso de los resultados de las evaluaciones: ¿castigar o ayudar?

“Desde fines de los 60 se ha medido el nivel de aprendizaje de los alumnos para elevarlo, pero no ha sido evidente la forma de hacerlo y, de hecho, no ha aumentado. La prueba de 8° grado (1967-1971) se limitó a mostrar a los profesores los tipos de habilidades que debían estimular en sus alumnos. La PER, creada a principios de los '80, buscó determinar el nivel del servicio educativo que se ofrecía y "monitorearlo" con ayuda de los padres, que tratarían de seleccionar las mejores escuelas para sus hijos. El SIMCE midió, a partir de 1988, los rendimientos de las escuelas y trató de identificar factores que pudieran explicar las diferencias y evaluar el impacto de los programas de los municipios y del Mineduc. Sin embargo, no hay avances y conviene repensar este poderoso instrumento, que es el SIMCE, si se quiere mejorar los aprendizajes. Un rol alternativo del SIMCE sería, según otros expertos, ayudar a los profesores a enseñar mejor y revisar cuáles son los temas importantes del currículo, más que el "clasificar" rendimientos. Este segundo rol plantea diferencias nítidas con respecto al primero: ¿Seleccionar o desarrollar? ¿Poner nota o medir logro de meta? ¿Conocer el ranking o el aprendizaje de objetivos específicos (criterio)? ¿Castigar o ayudar?

“Son dos funciones diferentes (los expertos hablan de Evaluación Sumativa y Formativa, respectivamente). Si la primera no elevó hasta ahora los rendimientos, convendría probar la segunda. Esto implica fuertes cambios en el SIMCE. En efecto, para "seleccionar" o "castigar" basta calcular un puntaje total que "discrimine entre buenos y malos" y difundirlo públicamente. En este rol es fundamental no divulgar los ítemes utilizados para que mantengan su poder de discriminación (no ser conocidos por los profesores y evitar que preparen a sus alumnos para contestarlos mecánicamente). Basta calcular el puntaje de cada escuela para que los padres seleccionen entre escuelas o preparar datos e indicadores globales para que los funcionarios del Mineduc modifiquen las estrategias.

“En la segunda función, en cambio, el usuario de la información del SIMCE es el profesor en su sala de clases. En efecto, para "desarrollar" o "ayudar a mejorar" se necesita entregar información detallada a cada profesor sobre los aspectos en que cada uno de sus alumnos logró los niveles adecuados y

aquellas habilidades o conocimientos que todavía no posee o no domina suficientemente. El profesor debe saber lo que contestó cada alumno en la prueba y revisar con cada uno los errores que cometió, hasta que el alumno internalice las deficiencias y reforme adecuadamente sus procesos de pensamiento. Sin información detallada el profesor no puede identificar los aspectos de su enseñanza que debe cambiar o cómo ayudar a cada estudiante...

“En resumen, una redefinición clara del objetivo de la prueba nacional de medición de la calidad de la educación debe orientar la reflexión del grupo de expertos sobre los cambios que debe tener el SIMCE...”

Fuente: Ernesto Schiefelbein, “SIMCE, ¿CASTIGAR O AYUDAR?”; publicado por el Diario La Tercera de Santiago de Chile, junio de 2003.

Un enfoque alternativo en el uso de los resultados: Just For The Kids

En el Estado de Tejas, en los Estados Unidos, funciona desde 1995 una organización no gubernamental y sin fines de lucro denominada “Just for the Kids”¹.

Su objetivo es colaborar en la mejora de la educación a través del uso intensivo de los resultados de las evaluaciones para mejorar la enseñanza.

Para ello, la organización publica en Internet los resultados de las escuelas en los exámenes estatales de una manera interesante, no bajo la forma de rankings, sino ofreciendo una visión de la evolución de cada escuela e introduciendo, como punto de comparación, el desempeño de los alumnos en las mejores escuelas de medios desfavorecidos.

Dicho en sus propias palabras, *“la idea es simple: ayudar a las escuelas a mejorar mostrándoles cómo otras escuelas con similar alumnado logran altos niveles de desempeño”*². Para ello, esta organización se dedica - además de a publicar la información en internet -- a la formación en servicio de maestros, a entrenar a los equipos de conducción escolar en relación a las “mejores prácticas” y a promover el compromiso público con el mejoramiento de las escuelas públicas. También realiza investigación.

¹ El sitio web de la organización es www.just4kids.org. El nombre de la organización juega con una doble acepción de la palabra “justo” como adjetivo vinculado al concepto de justicia o equidad, y como adverbio vinculado a las ideas de pertinencia, exactitud y/o exclusividad “sólo para los niños”.

² Véase JUST FOR THE KIDS, 2001; Promising Practices. How High-Performing Schools in Texas get Results. Austin, JUST FOR THE KIDS.

Un dato relevante que el lector debe tener en cuenta -y que hace posible esta experiencia- es que el Estado de Tejas tiene un sistema estatal de pruebas de tipo censal y criterial, es decir, que involucra a todas las escuelas y que establece categorías definidas en relación a cuáles son los niveles de desempeño aceptables para los alumnos. Estas pruebas se realizan todos los años.

Las figuras 1 y 2 ilustran el modo en que esta organización presenta en Internet los resultados de los centros educativos.

La figura 1 incluye, en la parte inferior de cada barra, los porcentajes de alumnos con resultado “proficiente” (satisfactorio) en la prueba estatal de Tejas (denominada TAAS).

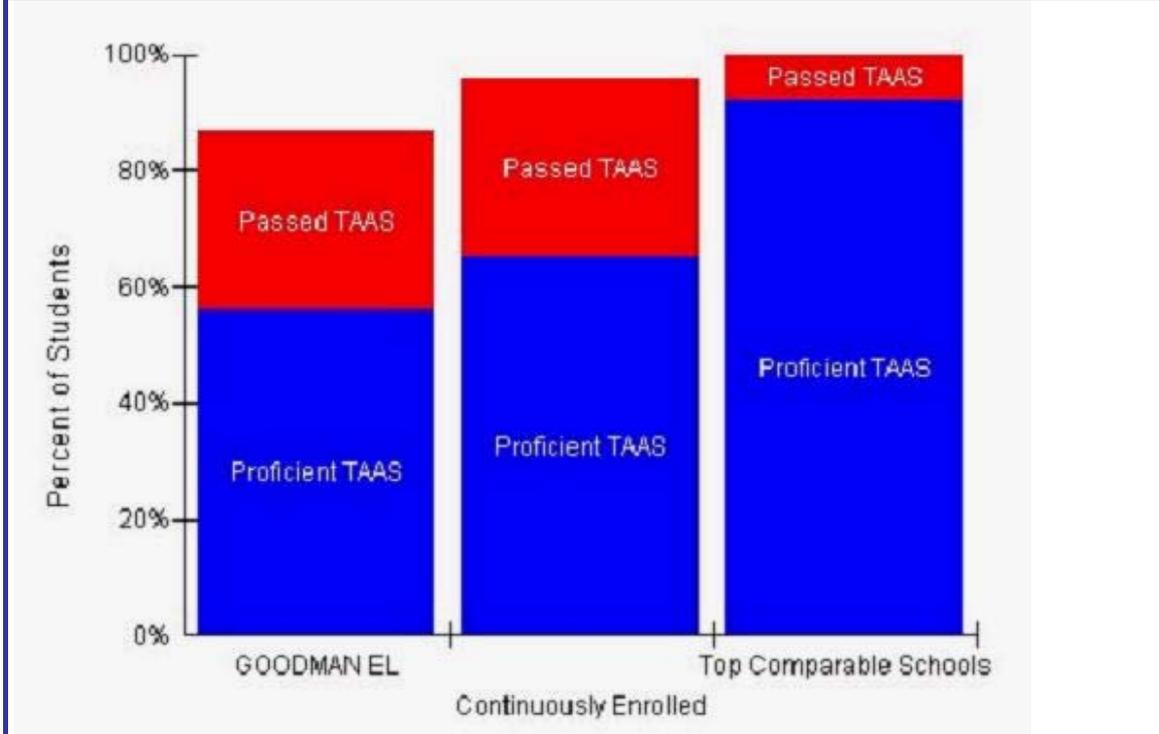
En la parte superior de cada barra se indica el porcentaje de alumnos que simplemente “pasaron” (aprobaron la prueba) pero sin alcanzar un nivel satisfactorio.

La primera barra del gráfico corresponde a una escuela en particular, la segunda incluye sólo a los alumnos que asistieron a la escuela en forma permanente, (definidos como aquellos que tenían por lo menos 2,8 años en la misma escuela), en tanto la tercera barra corresponde a los resultados de las mejores escuelas comparables con la escuela en cuestión.

Se define como “comparables” a aquellas escuelas que tienen igual o mayor proporción de alumnos provenientes de familias desfavorecidas o con dominio limitado del idioma inglés.

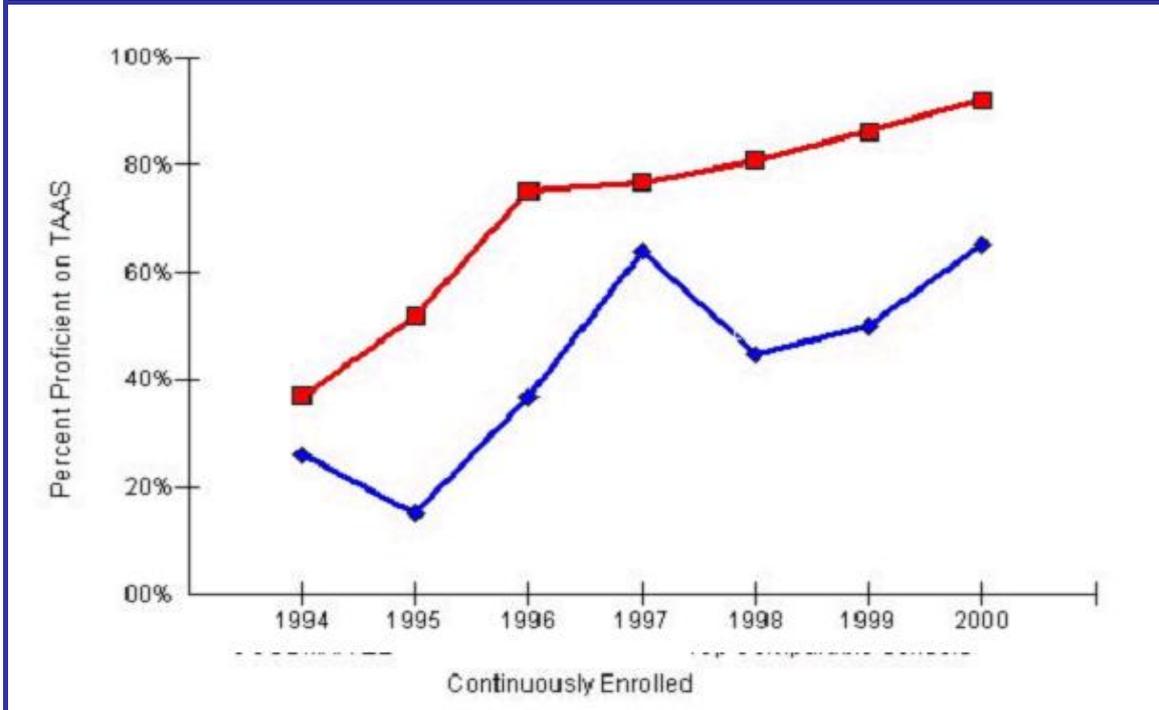
De este modo, la escuela y las familias pueden tener una idea del potencial de mejora de la escuela de sus hijos.

Figura 1
Los resultados de una escuela y las mejores "comparables"



Fuente: Just For The Kids, www.just4kids.org.

Figura 2
La evolución en el tiempo de los resultados de la escuela



Fuente: Just For The Kids, www.just4kids.org.

La figura 2 ofrece una mirada de la evolución a lo largo de los años (eje x) del porcentaje de alumnos que lograron un desempeño satisfactorio (eje y), en una escuela en particular (la línea de abajo, en azul) y en el conjunto de las escuelas "comparables" que obtuvieron los mejores resultados (la línea de arriba, en rojo). El gráfico considera únicamente a los alumnos matriculados en forma estable en la escuela.

Esto permite a familias, educadores y autoridades hacer seguimiento del impacto de las iniciativas y proyectos emprendidos por las escuelas. La información se entrega para cada grado y para cada asignatura.

También se publica en Internet una ficha de cada escuela y de las mejores, en las que se indican los datos demográficos básicos que deben ser tenidos en cuenta en las comparaciones: qué porcentaje de los alumnos pertenecen a hogares con bajos ingresos; qué porcentaje de los alumnos pertenecen a grupos de inmigrantes que tienen dominio limitado del inglés; qué porcentaje representan los alumnos que rindieron la prueba sobre el total de la matrícula; qué proporción de la matrícula total representan los alumnos que rindieron la prueba y son alumnos estables en la escuela.

Un trabajo de investigación realizado por Just For Kids sobre qué caracteriza a las escuelas que atienden poblaciones desfavorecidas (familias de bajos ingresos o pertenecientes a minorías étnicas) y tienen los mejores resultados, muestra que detrás de dichos resultados existen 5 elementos principales:

1. **Tomar la iniciativa: no dar excusas sino esforzarse por lograr que los alumnos aprendan.** Esto significa que las escuelas se proponen metas ambiciosas y creen que todos sus alumnos pueden aprender. Los directores de estas escuelas visitan continuamente las aulas y dialogan con los maestros en relación a cómo mejorar la enseñanza.
2. **Desarrollar y llevar adelante una estrategia clara para mejorar.** Se establece metas para cada grado y

asignatura y se crea un currículo coherente para toda la escuela; se adopta prácticas de enseñanza sustentadas en investigación empírica; se reestructura la jornada o el calendario escolar para que los docentes dispongan del tiempo necesario para planificar en forma conjunta y compartir y analizar sus prácticas; se le da a los maestros el apoyo, los materiales y la capacitación que necesitan.

3. **Evaluar continuamente el progreso de los alumnos e intervenir en forma inmediata cuando los alumnos o los maestros están en problemas.** Las evaluaciones en estas escuelas no son disparos al azar, sino que son cuidadosamente seleccionadas y aplicadas regularmente, tanto de manera formal como informal. Los maestros intervienen en forma inmediata para brindar apoyo específico a los alumnos que muestran problemas, incluyendo tutorías fuera del horario o calendario regular. También se interviene en forma inmediata cuando un maestro tiene problemas de orden conceptual o en el manejo de la clase.
4. **Hacer que la prioridad número uno sea la enseñanza de alta calidad y la práctica basada en la investigación.** El desarrollo profesional está centrado en las prioridades de enseñanza de la escuela y la capacitación del personal es la estrategia para lograr las metas de aprendizaje. Se contrata a especialistas para observar y orientar a los docentes. Los maestros nuevos son colocados al lado de un maestro experimentado para que lo apoye en la planificación de sus clases.
5. **Colaborar dentro y fuera de la escuela.** Los maestros trabajan en forma conjunta, como grupo de resolución de problemas. Los maestros de un mismo grado tienen momentos para planificar en común. El lenguaje escrito y la matemática son preocupación de los docentes de todas las áreas. Varias veces al año los docentes se reúnen por área o asignatura para planificar y coordinar en forma vertical (se refiere a los maestros

de grados sucesivos) el trabajo de toda la escuela.

Otra investigación, dirigida a analizar el impacto del uso de estos resultados por parte de las escuelas, muestra que las escuelas que realizaron al menos uno de los siguientes tres pasos: --estudiar escuelas de muy buenos resultados, establecer metas de aprendizaje claras, realizar cambios en el programa escolar o en el desarrollo profesional de los maestros--mejoraron entre 3 y 8 puntos porcentuales la "proficiencia" en lectura, escritura y matemática en los primeros grados¹.

La experiencia de Just For The Kids es ilustrativa de varias de las reflexiones formuladas en fichas anteriores, en particular de tres aspectos esenciales:

1. la importancia de generar y divulgar ampliamente información sobre los aprendizajes de los alumnos;

2. la importancia de hacerlo bajo un enfoque criterial;

3. la importancia de enfocar los esfuerzos hacia la creación de capacidades para enseñar, más que hacia los incentivos externos o la creación de un mercado competitivo.

Esto último incluye un aspecto de fundamental importancia que muestra la investigación realizada por esta organización. No basta con distribuir o publicar resultados; es imprescindible generar los espacios de trabajo e intercambio entre docentes para que los mismos sean analizados. La mejor publicación impresa nunca podrá sustituir el trabajo de análisis colectivo de los maestros. Si estos espacios no existen, difícilmente las evaluaciones contribuirán a mejorar la enseñanza.

Informando a familias y docentes: la experiencia del Territorio Capital de Australia (ACT)

El Territorio Capital de Australia tiene una interesante experiencia de evaluación, con un enfoque formativo (véase el recuadro 2). Dentro de este enfoque, dos modalidades de divulgación y uso de los resultados son particularmente interesantes. En primer lugar, la divulgación a los padres se realiza informándoles sobre el desempeño de su propio hijo, entregándoles una descripción detallada de los niveles de desempeño para cada competencia, al estilo de lo ilustrado anteriormente en las fichas 6 y 7. Las figuras 3.1 y 3.2 ilustran esta experiencia. Cada padre recibe algo así como un mapa de los desempeños en cada área evaluada y de la situación de su hijo.

Es importante poner de manifiesto la diferencia entre este enfoque y el de la información global sobre la escuela. Si bien no son enfoques incompatibles (de hecho, en el Territorio Capital de Australia estaba en estudio el modo de reportar a las familias resultados generales de la escuela en

relación al resto, pero sin establecer rankings, con un enfoque bastante similar al de Just For The Kids), informar a los padres acerca de qué se espera que sus hijos aprendan y en qué grado sus propios hijos lo están logrando permite comprometer a los padres con el aprendizaje de sus propios hijos, los ayuda a comprender qué es lo que se está intentando enseñar y abre las puertas para que ellos puedan colaborar con la escuela apoyando a su propio hijo o hija.

Las escuelas reciben un informe que compara los resultados de sus alumnos con los de otras escuelas, en tanto los docentes reciben información sobre el resultado de cada uno de sus alumnos, y como elemento de comparación, el porcentaje de alumnos que respondió correctamente cada pregunta en todo el Territorio Capital. Con estas informaciones, cada escuela debe elaborar un informe de análisis siguiendo la pauta incluida en el recuadro 3.

Recuadro 2

La experiencia del Territorio Capital de Australia

“El Territorio Capital de Australia empezó a aplicar pruebas de lenguaje y matemáticas en tercer, quinto, séptimo y noveno años en todas las escuelas gubernamentales entre 1997 y 1999. Las evaluaciones se basan en los resultados estipulados en los marcos del perfil curricular del ACT. Las capacidades de lectura, escritura y matemáticas son vistos como las piezas fundamentales de una educación exitosa, y como vitales para las oportunidades futuras de los estudiantes.

“El Programa está basado en un modelo referido a criterios y reporta la proporción de estudiantes que ha demostrado las capacidades documentadas en el perfil y, en tercer y quinto años, en los estándares nacionales.

“En el presupuesto de 1995, el gobierno del ACT anunció que se introducirían evaluaciones en dos grados de las escuelas primarias. Un grupo de referencia constituido por miembros del Foro de Consejos Escolares, el Consejo del ACT de Padres de Familia y Apoderados y de Asociaciones Ciudadanas, de la Asociación de Directores de Primaria, la Asociación de Directores de Secundaria, el Sindicato Educativo Australiano y del Departamento de Educación y Capacitación asesoró al gobierno con respecto a:

- el instrumento externo más apropiado disponible nacionalmente;
- los grados en los cuales se conducirían las evaluaciones;
- el formato para reportar datos a los padres y apoderados y al gobierno;
- arreglos para que los padres y apoderados pudieran optar por incluir o no a sus niños;
- el uso apropiado de los datos, tomando en consideración consideraciones de privacidad;
- establecer la relación entre reportar los resultados estudiantiles y los objetivos sociales de la escolaridad.

“El grupo de referencia examinó todos los instrumentos de evaluación disponibles en Australia que eran consistentes con los marcos curriculares del ACT y con el desarrollo curricular realizado en las escuelas.

“Se hizo un esfuerzo considerable por mantener una comunicación abierta con las escuelas y con la comunidad. Se mantuvo contacto regular con docentes, consejos escolares y organizaciones de padres de familia y ciudadanas para informarles sobre el programa y obtener su apoyo desde antes de su implementación. Se usaron notas de prensa y otros medios para informar a la comunidad en general y mantener actitudes positivas de la comunidad hacia el programa.

“El grupo de referencia consideró que la privacidad de los estudiantes, docentes y escuelas individuales era un factor crítico para el éxito del programa en un sistema tan pequeño. La publicación de los resultados de las escuelas podría dar información equívoca sobre la calidad de la enseñanza y los programas escolares. Los resultados de una escuela en particular están determinados por una gama de factores que no están bajo el control de escuelas o maestros individuales. En una jurisdicción pequeña, la publicación de los resultados de las escuelas puede desestabilizar las matrículas, lo que podría generar problemas de recursos. Intentos de comparar a estudiantes individuales podría ir en detrimento de su desarrollo social y académico. Por lo tanto, se decidió que los datos no se usarían para:

- comparar individuos;
- evaluar el desempeño docente;
- comparar a maestros individuales;
- hacer comparaciones positivas o negativas sobre el desempeño de las escuelas;
- comparar escuelas gubernamentales y no gubernamentales.

“La confidencialidad en relación a este asunto estaba cubierta por la Ley de Privacidad de 1988, y por los Estándares de Gestión Pública del ACT. Se reforzó esa confidencialidad mediante el desarrollo de un Protocolo de Confidencialidad que todo individuo que tuvo acceso a los datos tuvo que firmar. Los reportes sobre las escuelas son conocidos sólo por las mismas escuelas”.

Fuente: Margaret Forster, M., Masters, G. & Rowe, K., 2001; *Measuring Learning Outcomes: Options and Challenges in Evaluation and Performance Monitoring*. Material preparado para el Curso “Opciones Estratégicas para la Reforma Educativa” del Instituto del Banco Mundial. Australian Council for Educational Research (ACER) / World Bank Institute (WBI).

Recuadro 3

Pauta para el análisis de resultados en las escuelas

Anualmente, las escuelas tienen que elaborar respuestas a las siguientes preguntas y adjuntarlas a sus Planes de Lenguaje y Matemáticas.

1. En Lectura y Escritura, y en Sentido Numérico, de Datos y Sentido Espacial, ¿qué aspectos del currículo han tenido mayor impacto en el hecho de que su escuela haya tenido un desempeño promedio, por encima o por debajo del promedio del sistema en algunos ítems específicos de la evaluación?

¿Cómo está la escuela preocupándose por el desempeño escolar en

- el nivel de grado

-el nivel de toda la escuela?

2. Si observa a los estudiantes de su escuela que están en el 20% inferior del ACT, ¿cuántos estudiantes hay en el 5%, 10% y 15% inferior?

Describa las intervenciones planificadas que han sido identificadas para mejorar el desempeño de estos estudiantes.

¿Cómo evaluarán los resultados de estas intervenciones, antes de la siguiente evaluación sistémica de los estudiantes?

3. De acuerdo con los datos de 1999, ¿cuánto crecimiento/declinación han experimentado los alumnos de quinto año desde que estaban en el tercer año?

¿Qué factores han causado este progreso o declinación?

4. ¿Cuán importante es la influencia del desempeño de los estudiantes de subgrupos por género, origen lingüístico no inglés u origen indígena sobre los resultados de su escuela?

Especifique las intervenciones que se están realizando para mejorar los desempeños de cada subgrupo (v.g., Programas Individuales de Aprendizaje para Estudiantes Indígenas).

Fuente: Margaret Forster, M., Masters, G. & Rowe, K., 2001; *Measuring Learning Outcomes: Options and Challenges in Evaluation and Performance Monitoring*. Material preparado para el Curso "Opciones Estratégicas para la Reforma Educativa" del Instituto del Banco Mundial. Australian Council for Educational Research (ACER) / World Bank Institute (WBI).

Figura 3.1. - Reportando a las familias (I)



Territorio de la Capital Australiana

Programa de Evaluación
Informe estudiantil quinto grado, 1999

Estimados padres de familia o apoderados:

Su niño rindió recientemente algunas pruebas de lenguaje y matemáticas, junto con otros estudiantes de quinto año de las escuelas públicas del ACT, utilizando los instrumentos desarrollados por ACER denominados DART. Las pruebas han sido calificadas externamente por ACER.

Evaluación en capacidades lingüísticas
Las tareas de lectura exigían que cada niño leyera cartas, textos informativos e ilustrados y un poema, y luego responder preguntas sobre lo que habían leído. Las tareas de lectura de imágenes requerían que los niños vieran un video y respondieran algunas preguntas por escrito. Las tareas principales de Escritura pedían a los estudiantes escribir una historia o cuento y un argumento. Antes de responder por escrito a las preguntas sobre capacidad de escucha, los estudiantes oyeron una cinta de audio, mientras que sus propios maestros evaluaron presentaciones orales para medir su capacidad en lenguaje oral.

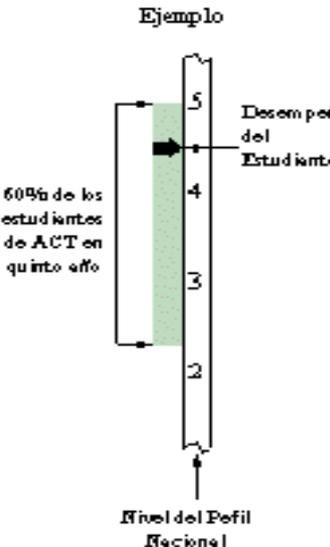
Evaluación en capacidades numéricas
Las tareas de sentido numérico exigían que cada niño manipulara y realizara cálculos con números presentados en tablas o recetas. Para algunas tareas los estudiantes podían usar una calculadora. Las tareas de sentido espacial involucraban interpretar mapas, fotos y diagramas, para demostrar su comprensión de conceptos espaciales tales como locación, forma, movimiento y orientación. Las tareas de medición y comprensión de datos implicaban medir y estimar tiempo, longitud, masa y área, y procesar e interpretar datos. Todas estas tareas estuvieron vinculadas al tema de Matemáticas en el Zoológico.

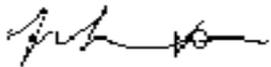
Estos resultados colocan a su niño en uno de cuatro niveles de logros de los ocho niveles del Perfil Nacional del Currículo de Inglés (para kindergarten a décimo año) y del Perfil Nacional del Currículo de Matemáticas (para kindergarten a décimo año) que describen las capacidades lingüísticas y numéricas que típicamente muestran los estudiantes en ese nivel.

Las áreas sombreadas muestran el logro del 60% de los estudiantes de ACT en quinto año.

Los resultados de su niño están marcados con una ➡ para cada aspecto de lenguaje y matemáticas que ha sido evaluado. Se puede ver un ejemplo adyacente a este texto.

Es importante leer este reporte junto con otra información proporcionada por el docente de aula, ya que hay otros aspectos escolares que son de significancia para el éxito de su niño en la escuela.





Fran Hinton
Oficial Principal

ACER Preparado por el Consejo Australiano de Investigación Educativa para el Gobierno de la ACT en 1999

Fuente: Margaret Forster, M., Masters, G. & Rowe, K., 2001; Measuring Learning Outcomes: Options and Challenges in Evaluation and Performance Monitoring. Material preparado para el Curso "Opciones Estratégicas para la Reforma Educativa" del Instituto del Banco Mundial. Australian Council for Educational Research (ACER) / World Bank Institute (WBI).

Figura 3.2. - Reportando a las familias (II)

CAPACIDADES DE LENGUAJE QUINTO AÑO 1999		
Lectura	Visualización	Escritura: contenido
<ul style="list-style-type: none"> Identifica el significado cultural de una imagen. Reconoce la relación entre un estilo de presentación y la naturaleza de la información (v.g., el formato de pregunta-respuesta en datos de entrevistas). Reconoce varias piezas de información refraseada de un texto largo. Infiere significado de lenguaje figurativo. Interpreta una relación que no ha sido hecha explícita. Ordena eventos detallados de una narrativa. Reconoce el tono de un poema simple. Hace conexiones entre ilustraciones y texto escrito. Selecciona letras con la misma idea clave. Interpreta la conexión entre una ilustración y un texto. Selecciona varias piezas de información de una presentación compleja de texto (busca ejemplos de información solicitada). Encuentra evidencia para sustentar una declaración. Localiza información en una carta. Reconoce rasgos lingüísticos convencionales (v.g. guías de pronunciación). Localiza información en una presentación compleja de texto. Conecta letras y materia temática en un formato de tabla. Ofrece una interpretación plausible de una porción de texto. Interpreta información fáctica. Interpreta una palabra inventada con referencia a una ilustración. Identifica la idea central en una carta (a partir de una lista dada). Reconoce la relación entre dos porciones de texto (v.g., interpreta el propósito de anotaciones). Hace conexiones entre palabras con significado similares (v.g., "ilustrador" traza dibujos). Infiere información de claves obvias en una porción simple de texto. Infiere el significado de jerga técnica a partir de su contexto escrito y visual. Localiza información en una presentación compleja de texto. Extrae información de una presentación compleja de texto y fotografía. Identifica el propósito de un texto escrito o visual computarizado. Identifica al autor de una carta a partir de su estilo contenido (v.g., una carta escrita por un niño). Infiere información implicada por el uso de palabra clave en un texto computarizado. 	<ul style="list-style-type: none"> Explica la consistencia de una escena final con referencia al tono de una película. Explica el detalle de un texto en términos de su contribución a la estructura del mismo (v.g., la música con que se inicia y termina una película). Identifica los elementos de una película que contribuyen a su tono (v.g., reconoce el uso de una parodia). Reconoce la audiencia objetivo de publicidad. Reconoce que las voces superpuestas pueden usarse para dar una perspectiva particular. Identifica una gama de técnicas que se usan para establecer un clima (v.g., ángulos de cámara, cámara lenta, música). Interpreta literalmente las palabras usadas al final de una película. Justifica su propia interpretación de un texto (v.g., puede escoger referirse a personaje, trama o tono cuando interpreta la selección de recursos cinematográficos). Deriva inferencias sobre ideas implícitas en un texto. Predice las instrucciones verbales para un procedimiento filmado, reconociendo el estilo instruccional. Resume una narrativa focalizándose en el tema central. Explica el detalle de un texto refiriéndose solamente a su contexto inmediato. Explica la escena final de una película con referencia al argumento o trama. Explica el significado central de un recurso cinematográfico. Explica el propósito general de la publicidad. Explica un vínculo simple entre acción visual y texto oral. Identifica y explica la selección de una localización para una película. Puede entender información contradictoria. Identifica la posición de la cámara en una toma. Entiende el propósito de una técnica cinematográfica básica. Construye argumentos razonables para distintos puntos de vista (v.g., fines alternativos de películas) Reconoce que las voces superpuestas se usan como recurso narrativo. Recuerda detalles de una película. Explica la escena final de una película sólo en términos de detalles periféricos. Recuerda algunos pasos en un procedimiento. Infiere los motivos de un personaje. Identifica un elemento clave de una secuencia narrativa. Resume una narrativa focalizándose solamente en detalles periféricos. Predice instrucciones verbales básicas para un procedimiento filmado. Identifica un recurso cinematográfico central. 	<p>Logros típicos de los estudiantes en cada nivel</p> <ul style="list-style-type: none"> Nivel 5 <ul style="list-style-type: none"> Escribe una narrativa desarrollada e integrada que incorpora el estímulo de manera convincente. Presta atención al ordenamiento temporal, un punto de vista narrativo consistente y el desarrollo de los personajes. Escribe un argumento coherente con evidencia convincente, aunque todavía no un caso totalmente desarrollado. Puede referirse a un contra-argumento. Involucra y persuade a su audiencia. Nivel 4 <ul style="list-style-type: none"> Escribe una narrativa bien construida en términos generales que incorpora bien el estímulo y tiene una sólida tipificación de los personajes. Escribe un argumento coherente, focalizado en la cuestión y con algún razonamiento, quizás incorporando un reconocimiento de una posición opuesta. Intenta involucrar o persuadir al lector. Nivel 3 <ul style="list-style-type: none"> Mantiene el foco de la narrativa, con eventos y detalles que contribuyen a la trama de la historia y con personajes identificables. Usa el estímulo para la historia de manera plausible en el propio mundo de ficción. Escribe un argumento basado en el estímulo, con un punto de vista claro y cierta distancia crítica, y puede referirse brevemente a ambos lados. Considera el impacto sobre la audiencia y puede imitar modelos publicados. Nivel 2 <ul style="list-style-type: none"> Muestra comprensión del género narrativo y habilidades emergentes en lo que se refiere a construcción de trama, con los personajes emergiendo a través del discurso o la acción. Da un punto de vista personal y general sobre una cuestión, con varias ideas como sustento, y puede identificar "el otro lado". Usa el estímulo para estructurar la redacción. Muestra alguna conciencia de la audiencia. Nivel 1 <ul style="list-style-type: none"> Escribe una narrativa con un flujo de historia reconocible, con estructura narrativa y con una localización clara, pero con definición mínima de personajes. Escribe una opinión basada en involucramiento personal, con unas cuantas ideas relacionadas. Vincula claramente un cuento o una opinión con el estímulo ofrecido. Escribe con alguna coherencia, que podrá no ser sostenida. Muestra un sentido emergente de audiencia. Incluye elementos de un cuento (puede ser corto), pero la narrativa carece de coherencia. Da una declaración simple de opinión, con alguna explicación, pero se apoya más en aseveraciones que en argumentos. Escribe un cuento o una opinión que tiene sentido para el lector. Nivel 0 <ul style="list-style-type: none"> Muestra alguna conciencia sobre la tarea. Escribe una respuesta muy corta o una larga y desconexa. No tiene control sobre la extensión de la historia, de la trama ni de los personajes. Escribe una historia personal en lugar de una opinión sobre un tema. Muestra poca, si alguna, conciencia de la audiencia.

Fuente: Margaret Forster, M., Masters, G. & Rowe, K., 2001; Measuring Learning Outcomes: Options and Challenges in Evaluation and Performance Monitoring. Material preparado para el Curso "Opciones Estratégicas para la Reforma Educativa" del Instituto del Banco Mundial. Australian Council for Educational Research (ACER) / World Bank Institute (WBI).

Aportando a inspectores y maestros herramientas para evaluar

Uruguay también ha desarrollado su experiencia de evaluación a nivel nacional desde 1995 con un enfoque de tipo formativo y sin consecuencias para las escuelas.

Un primer aspecto a destacar de esta experiencia es un modo de comparar resultados entre escuelas que evita los problemas analizados en la ficha 10 relativos

a los rankings. En este país, cuando las evaluaciones tienen carácter censal, cada Inspección Departamental recibe un "mapa" de sus escuelas, como el incluido en la figura 4.

En cada casilla se incluye grupos de escuelas, clasificadas en función de su contexto sociocultural y de sus resultados.

Figura 4
"Mapa socioacadémico" distribuido a los cuerpos inspectivos

Resultado en la prueba de Lengua	Contexto sociocultural de la escuela				
	Muy Favorable	Favorable	Medio	Desfavorable	Muy Desfavorable
Alto 75,1% a 100% de alumnos suficientes	Escuelas:	Escuelas:	Escuelas:	Escuelas:	Escuelas:
Medio-alto 57,5% a 75% de alumnos suficientes	Escuelas:	Escuelas:	Escuelas:	Escuelas:	Escuelas:
Medio-bajo 39,4% a 57,4% de alumnos suficientes	Escuelas:	Escuelas:	Escuelas:	Escuelas:	Escuelas:
Bajo 0% a 39,3% de alumnos suficientes	Escuelas:	Escuelas:	Escuelas:	Escuelas:	Escuelas:

Fuente: ANEP, Unidad de Medición de Resultados Educativos, 2000; Evaluaciones Nacionales de Aprendizajes en Educación Primaria en el Uruguay. 1995-1999. UMRE, Montevideo.

Este modo de organizar la información sobre las escuelas tiene varias ventajas:

- a. Evita ofrecer un ordenamiento posiblemente erróneo de las escuelas, como ocurre en el caso de los rankings (véase la ficha 10), proponiendo una clasificación de las escuelas en cuatro grandes grupos en función de los resultados de sus alumnos.
- b. Estos grupos no están definidos a partir de promedios, sino a partir del porcentaje de alumnos que alcanzaron un nivel de suficiencia en la prueba de

Lengua. En este caso, se considera como escuelas con alto rendimiento a aquellas en que más de 3 cuartas partes de su alumnado alcanzaron dicho nivel. El grupo medio-alto está conformado por las escuelas cuyo porcentaje de alumnos suficientes estuvo entre la media nacional (57,5%) y el 75%.

- c. Los grupos se conformaron por separado, en función del contexto sociocultural de la escuela. De este modo se evitó la comparación entre escuelas con poblaciones muy diferentes. Al mismo tiempo, es posible identificar a las

escuelas con muy buenos resultados en contextos desfavorecidos (la zona del "mapa" que aparece en gris).

Uno de los aspectos más interesantes de la experiencia de Uruguay en materia de evaluaciones estandarizadas ha sido la denominada "evaluación muestral con aplicación autónoma en el conjunto de las escuelas".

Esta estrategia busca mantener el involucramiento de todas las escuelas en la evaluación nacional, sin incurrir en el costo de una evaluación censal. Lo que se hace es evaluar en forma controlada desde la unidad central una muestra representativa pero pequeña de escuelas, para luego distribuir las pruebas al conjunto de todas las escuelas del país, junto con manuales para la aplicación y corrección de las pruebas en forma autónoma por parte de las escuelas.

Cada escuela participa en forma voluntaria y los resultados de su aplicación son manejados únicamente a nivel de la escuela. La escuela recibe luego los resultados nacionales y por contexto sociocultural, con una columna en blanco para ubicar allí los resultados de la propia escuela, como forma de compararse con los resultados nacionales y con los resultados de escuelas de similar población.

A modo de ejemplo, todas las escuelas del país recibieron las pruebas para los niños de 5 años, 1° y 2° grados (aquellas que utilizaron los niveles de desempeño presentados en el ejemplo 5 de la ficha 7), de modo que pudieron aplicarlas a sus alumnos y compararse con los resultados nacionales.

La experiencia ha mostrado que los maestros encuentran mucho más enriquecedor este ejercicio de evaluación autónoma que cualquier resultado nacional de una prueba en la que no participan directamente. De acuerdo a la investigación realizada en el marco de PREAL (Kaztmann, R., Aristimuño, A. B. Monteiro, L., 2003; *¿Cómo se usa y qué impacto tiene la información empírica sobre las evaluaciones nacionales de aprendizajes en el mejoramiento de la Educación Primaria pública en Uruguay?*; PREAL/GDN) en una encuesta representativa a nivel nacional alrededor del 80% de los maestros declara que participan de la aplicación autónoma de las pruebas.

Otra ventaja de este enfoque es que permite trabajar con pruebas de respuesta construida en lugar de pruebas de elección múltiple, dado que la muestra que debe ser corregida centralmente es relativamente pequeña en comparación con un censo.

Información para las autoridades y la opinión pública: un desafío complejo

Si bien su tratamiento exhaustivo excede los límites de esta ficha, es necesario mencionar otras dos audiencias clave de los resultados de las evaluaciones, además de los educadores y las familias: las autoridades y la opinión pública -- en este último caso, a través de los medios de prensa.

En relación a las autoridades, es preciso señalar la importancia de que las unidades de evaluación desarrollen su capacidad para entregar a las autoridades resúmenes ejecutivos que presenten con claridad lo principal de la información producida, así como análisis de situación que den una visión más amplia de los problemas educativos que la evaluación pone de manifiesto.

Esto último seguramente requiere de la discusión abierta de los resultados con diversos actores, dado que de los resultados no surge automáticamente el diagnóstico ni los lineamientos de trabajo. Es necesario un trabajo de investigación y de reflexión y análisis, que necesariamente debe involucrar a diversos actores.

Para las autoridades es importante, además, la información relativa a los factores que tienen incidencia sobre los resultados. Este aspecto es tratado en la ficha 13.

En relación a la opinión pública, el problema central es que normalmente ni las unidades de evaluación ni los ministerios cuentan con recursos para desarrollar una campaña de tipo informativo, razón por la cual el modo en que los resultados se divulgan queda

librado a los enfoques que la prensa quiera dar. Y, lamentablemente, la experiencia demuestra que, en general, predomina en la prensa un abordaje de tipo sensacionalista, que busca generar escándalo como modo de vender, más que la contribución a la información y al debate ciudadano.

Este aspecto ha sido investigado en el informe *¿Cómo aparecen los resultados de*

las evaluaciones educativas en la prensa?, disponible en el sitio web del Grupo de Trabajo de Estándares y Evaluación de PREAL (www.preal.cl/gtee).

En ese trabajo se analizan los tratamientos que la prensa da a los resultados de las evaluaciones, los errores más comunes que se cometen y recomendaciones para mejorar la forma de reportar.

Síntesis final

A lo largo de esta ficha se intentó ilustrar con ejemplos específicos cómo las evaluaciones estandarizadas pueden ser utilizadas como herramientas, para el aprendizaje y para ayudar a las escuelas y a los docentes a mejorar su trabajo, desde un enfoque formativo antes que punitivo.

Las evaluaciones estandarizadas, sus instrumentos y sus resultados pueden ser utilizados para mejorar el modo en que los maestros evalúan y para identificar y dar a conocer experiencias de escuelas con niveles de logro destacados en contextos difíciles.

Para ello es fundamental producir estrategias y materiales de difusión específicamente pensados para apoyar el trabajo docente y, en especial, articular la divulgación de resultados con programas de capacitación en servicio. En este sentido, la articulación entre las unidades de evaluación y las áreas, programas e instituciones que desarrollan programas de formación y capacitación de docentes es clave para propiciar la reflexión a partir de los resultados de las evaluaciones y el cambio de las prácticas de enseñanza.

Pueden ser utilizados también para informar, involucrar y comprometer a las familias con el aprendizaje de sus hijos, pero desde la perspectiva de construir una relación de apoyo y colaboración con la escuela en una tarea común --educar a los niños y jóvenes--, en lugar de generar una situación de enfrentamiento o una relación "cliente-proveedor".

La experiencia internacional indica que este tipo de aproximaciones favorecen mejor el desarrollo de una cultura de la evaluación y actitudes positivas hacia las evaluaciones externas que cuando las mismas tienen un carácter únicamente de control o una connotación punitiva.

ⁱ Dougherty, C. & Collins, S. (2002); Use of the Just for the Kids Data By Texas Elementary Schools. Austin, Just for The Kids.