

## VALOR AGREGADO EN EDUCACIÓN: QUÉ ES Y CÓMO SE USA

El presente número del Boletín del Observatorio de Políticas de Evaluación Educativa tiene como propósito informar acerca del uso de medidas de valor agregado para valorar el trabajo de escuelas y docentes. Dicho uso es creciente en los Estados Unidos y en el Reino Unido, países que cuentan con grandes bases de información sobre el desempeño de sus estudiantes y sobre diversidad de variables relacionadas con su entorno, sus escuelas y sus docentes.

A lo largo del boletín se intenta explicar cuestiones básicas para comprender el concepto de valor agregado, ejemplificar buenos y malos usos del mismo, e informar al lector acerca de los cuidados que deben tenerse al recurrir a este tipo de medida.

Dado su origen, las figuras y ejemplos están en inglés, al igual que los sitios web de origen. Ello puede dar lugar a alguna dificultad a la hora de comprenderlas, si bien se ha intentado explicar su contenido.

### 1. Qué es el valor agregado

El valor agregado (VA) es un concepto utilizado en la investigación educativa, con el fin de medir el progreso de los alumnos. Es un modelo para el análisis de datos de aprendizaje. La posibilidad de aplicarlo depende del diseño utilizado en cada estudio. En la última década está siendo cada vez más utilizado en los Estados Unidos para la medición del aporte de las escuelas y los docentes al aprendizaje de sus alumnos, en el marco de programas de incentivos por resultados, responsabilidad docente y mejora de las escuelas.

A continuación se presentan dos definiciones del concepto, complementarias entre sí:

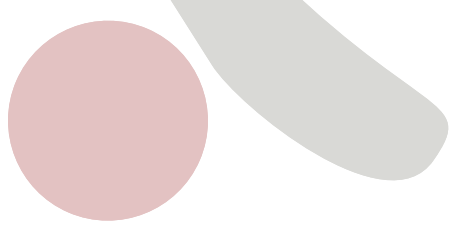
“Los modelos de VA son un conjunto de procedimientos estadísticos que se utilizan para hacer inferencias sobre la eficacia de las escuelas y de los profesores que ponen el acento en las ganancias de los estudiantes en el tiempo. Tienen en común el seguimiento de la trayectoria de los estudiantes analizando las medidas de los resultados de dos o más años. Estos datos de la evolución de los estudiantes se transforman en indicadores de la eficacia de la escuela o del profesor” (Martínez Arias, Gaviria y Morera, 2009:17).

“En su forma más básica, la evaluación del valor agregado se basa simplemente en el cálculo del cambio año a año en los puntajes de los estudiantes; otras formas más complicadas incorporan técnicas estadísticas para tener en cuenta factores tales como las diferencias en las características de los estudiantes y efectos persistentes de anteriores docentes y escuelas” (Wiley, s/f: 5. Traducción propia).

Podemos decir entonces que el valor agregado es un modelo para analizar la trayectoria escolar de un mismo conjunto de niños en cierto período de tiempo, por oposición a las mediciones que se realizan en un único momento. El foco en la medición del progreso de los niños permite controlar el efecto de variables externas a la experiencia escolar y las diferencias en los puntos de partida de los niños. Por ello es una herramienta más apropiada para conocer cuánto agregan escuelas y/o docentes al aprendizaje esperado de los alumnos. El valor esperado para cada alumno se define en función de su desempeño previo, para lo cual pueden o no incorporarse algunas características propias de los alumnos, profesores y escuelas. La diferencia entre el desempeño final real y el esperado es el valor agregado.

El uso de modelos de valor agregado intenta corregir la principal debilidad de los rankings de escuelas o maestros generados a partir de una única medición de logros educativos. En este último caso, normalmente los rankings reflejan más el origen social de los estudiantes que la calidad de su experiencia escolar.<sup>1</sup>

<sup>1</sup> Véase el Boletín N° 16 de esta serie, sobre el caso de Chile.



## 2. Diferentes usos del valor agregado

Los modelos de valor agregado pueden utilizarse en cualquier evaluación educativa que releve información sobre los desempeños de los mismos individuos en dos momentos del tiempo. Sus usos más extendidos son:

- en investigación educativa, cuando se busca identificar, por ejemplo, factores del sistema que inciden en el aprendizaje de los alumnos;
- en la evaluación del impacto de programas educativos en los aprendizajes de los estudiantes;
- en programas de mejora de las escuelas;
- en sistemas de rendición de cuentas que buscan identificar el efecto de los maestros y/o efecto de las escuelas en la mejora de los aprendizajes de sus alumnos.

En la actualidad, los usos que resultan más controversiales son los orientados a la mejora escolar y a la rendición de cuentas.

El modelo aplicado actualmente en Inglaterra es un ejemplo de uso orientado a la mejora escolar. Dicho modelo tuvo un impulso importante a partir de la ley *Every Child Matters*<sup>2</sup>, que si bien en un principio estuvo orientada hacia la rendición de cuentas luego fue cambiando su orientación: “La práctica de la Inspección ha pasado del foco en la rendición de cuentas a estar más enfocada en el apoyo para la mejora” (PREAL/GTEE, 2010:2). Este cambio requirió una modificación en la forma de presentación de los resultados de las mediciones externas: se pasó de tablas de logro educativo, que incluían simples rankings de escuelas, a los actuales reportes de valor agregado, contextualizado por distintos factores.

En Estados Unidos, la mayor parte de los modelos de valor agregado está orientada hacia sistemas de rendición de cuentas. Las experiencias son muy variadas, desde mediciones de valor agregado a nivel de escuelas hasta mediciones de valor agregado centradas en cada docente en particular. En muchos distritos se toman decisiones sobre los salarios de cada docente a partir de su valor agregado.

En estos casos, dichos modelos:

- Consisten en “cómo medir el desempeño de los alumnos en las escuelas del país en un sentido que arroje luz en el progreso, así como en los estándares” (Saunders, L. 2000:241. Traducción propia).
- Se usan “para estimar el efecto verdadero de un docente o escuela en particular sobre el aprendizaje de los estudiantes” (Wiley, s/f: 2. Traducción propia).
- Son “un conjunto de técnicas estadísticas complejas que usan datos de puntajes de pruebas de los estudiantes en múltiples años para estimar los efectos de escuelas o profesores individuales” (McCaffrey et al, 2003: xi. Traducción propia).

Al respecto, el Brown Center on Education Policy indica que “el valor agregado tiene un rol importante a jugar en los sistemas de evaluación docente, pero todavía hay mucho por aprender sobre cómo usar mejor la información de valor agregado en las decisiones en recursos humanos” (2010:1. Traducción propia).

Partiendo de la idea central de que “los datos sobre resultados de los alumnos pueden ser usados como parte de la evaluación docente” (Sanders, 2000:329. Traducción propia), los modos de hacerlo son variados (Goe, 2008):

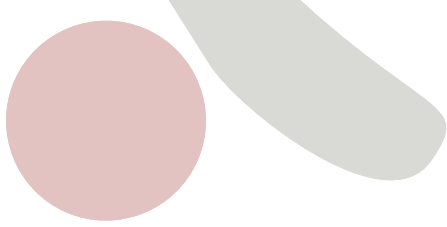
- a. Como forma de identificar profesores efectivos e inefectivos.
- b. Como parte de la evaluación del desempeño docente.
- c. Como ayuda para tomar decisiones sobre salarios diferenciados entre docentes.
- d. Para la certificación docente.
- e. Para mejorar el desarrollo profesional.
- f. Para mejorar los programas de formación docente.

Según Wright, Sanders y Rivers, estos modelos también pueden ser utilizados para el seguimiento de alumnos, ya que sus “proyecciones proveen una vía conveniente para identificar si un alumno está o no en camino a ser competente” (2006:11. Traducción propia). Los autores señalan que este enfoque contribuye a alcanzar los objetivos de la ley *No Child Left Behind*<sup>3</sup>, según la cual “en escuelas que no alcanzan los estándares estatales al menos en dos años consecutivos, los padres podrán transferir a sus hijos a una escuela pública con mejores resultados (...) Los alumnos provenientes de familias de bajo nivel socioeconómico que no alcancen los estándares estatales por al menos tres años son elegibles para recibir servicios educativos suplementarios, incluyendo tutorías, servicios en el contra horario y escuela de verano”<sup>4</sup>.

<sup>2</sup> <http://www.education.gov.uk/consultations/downloadableDocs/EveryChildMatters.pdf>

<sup>3</sup> <http://www2.ed.gov/nclb/>

<sup>4</sup> Fuente: <http://www2.ed.gov/nclb/overview/intro/4pillars.html>. Traducción propia.



En particular sobre la evaluación docente, el National Comprehensive Center for Teacher Quality (Centro Nacional Integral para la Calidad Docente) dice que:

“En su forma más simple, la medida de valor agregado, tal como es usada para la evaluación de docentes, se calcula como sigue: los puntajes iniciales de los estudiantes son usados para crear puntajes esperados en un test para cierto año. La diferencia entre el puntaje esperado o predicho y el puntaje realmente obtenido en ese año, es el crecimiento. La contribución de los docentes al aprendizaje de los alumnos se determina observando el promedio del crecimiento en el puntaje de todos los alumnos de cada docente. Los docentes son ranqueados en relación a otros docentes de un mismo distrito (u otra unidad de interés) de acuerdo a cuanto cada uno contribuyó al crecimiento del puntaje de los estudiantes, y este ranking es su 'puntaje' de valor agregado. En algunos modelos de valor agregado, sólo se usan en el cálculo los puntajes previos de los estudiantes, otros modelos incluyen el sexo, raza y nivel socioeconómico, otros incluyen información acerca de la experiencia docente. Con una medida de valor agregado, los docentes cuyos estudiantes obtuvieron un resultado tan bueno como el que se predijo son considerados “encima del promedio” o “altamente efectivos”, y aquellos cuyos estudiantes obtuvieron resultados peores que los esperados son considerados “debajo del promedio” (Goe, L., 2008:2. Traducción propia).

En un trabajo de la Universidad de Colorado, Wiley realiza una descripción de los modelos de valor agregado, de los cuales tres han sido utilizados en Estados Unidos para rendición de cuentas (s/f:2):

- Simple gain score models (ej. Texas Growth Index).
- Covariate adjustment models (ej. Dallas Value Added Assessment System or DVAAS).
- William Sanders's layered models (ej. Educational Value Added Assessment System or EVAAS; Tennessee Value Added Assessment System or TVAAS).
- Cross-classified models.
- RAND's persistence model.
- Todd and Wolpin's cumulative within-child mixed-effects model.

Por otra parte, Martínez Arias describe “los modelos con mayor implantación en la práctica”, entre los que incluye (2009:220.222):

- Tennessee Value Added Assessment System EVAAS o TVAAS
- Dallas Value Added Assessment System DVAAS
- Valor añadido de las escuelas públicas de Chicago
- El modelo de Valor Añadido Contextualizado en Inglaterra
- Algunas experiencias españolas

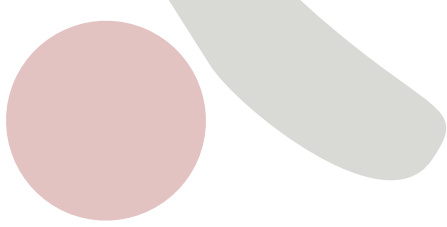
El más conocido de estos es el desarrollado por William Sanders en Tennessee (TVAAS/EVAAS<sup>5</sup>), por haber sido implementado tempranamente en un estado de Estados Unidos (1992). Este modelo es actualmente el sistema de evaluación de valor agregado más utilizado en Estados Unidos.<sup>6</sup>

### **3.1. Ejemplos de usos del valor agregado. Qué información se brinda en la evaluación de escuelas y docentes.**

En este apartado se presentan algunos ejemplos ilustrativos de las formas en que se presentan los resultados a las escuelas en Tennessee y en Inglaterra. Luego se muestra un ejemplo del uso del valor agregado en evaluación docente en Estados Unidos.

<sup>5</sup> Tennessee value added assessment system.

<sup>6</sup> [http://www.cgp.upenn.edu/ope\\_tn.html](http://www.cgp.upenn.edu/ope_tn.html) y [http://www.cgp.upenn.edu/ope\\_value.html](http://www.cgp.upenn.edu/ope_value.html)



### 3.1. Valor agregado en Tennessee

En el estado de Tennessee los resultados sobre valor agregado pueden consultarse a nivel de estado, sistema o escuela. Al seleccionar una escuela, la información disponible es:

- un perfil de la escuela en donde se indica qué grados atiende, la distribución de los alumnos según raza y sexo, así como la proporción en condición económica desfavorecida,
- la clasificación de la escuela según los criterios de No Child Left Behind desde el año 2005,
- los puntajes en distintas áreas y grados,
- para cada área y grado, la ganancia en el puntaje por año, el promedio de todos los años considerados, comparativamente con la ganancia del estado,
- el porcentaje de asistencia, promoción y abandono escolar,
- la proporción de alumnos que fueron suspendidos y expulsados, y
- la proporción de cursos dictados por profesores altamente calificados.<sup>7</sup>

La imagen que sigue muestra el formato de presentación de resultados recién descrito. En esta ventana aparece la opción de solicitar los resultados según estado, sistema o escuela y más abajo se ven distintas pestañas con la información disponible. En este caso se seleccionó la escuela Anderson Elementary del sistema Bristol Schools.

The screenshot shows the Tennessee Department of Education website. At the top, it says 'TN GOV Department of Education Dr. Tim Webb, Link back to Home.' Below that, there are navigation tabs for 'TDOE Report Card', 'Report Card Explained', 'Report Card Terms', 'Benchmarks & Documents', and 'NAEP data'. The main content area displays 'TENNESSEE DEPARTMENT OF EDUCATION Report Card 2009' and 'BRISTOL SCHOOLS Director: Dr. Steve Dixon ANDERSON ELEMENTARY Principal: Mr. Andrew Brown'. There is a dropdown menu for 'Report Card for 2009' and radio buttons for 'State', 'System', and 'School'. The 'School' option is selected, and 'Anderson Elementary' is chosen. At the bottom, there are tabs for 'Profile', 'NCLB (AYP)', 'Achievement', 'Value-Added', 'Attendance and Graduation', 'Discipline', and 'Teacher'. The 'Value-Added' tab is currently active.

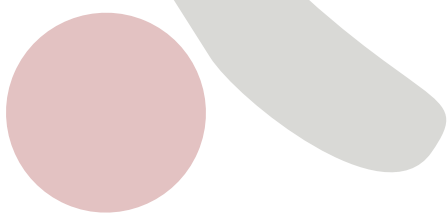
En la pestaña valor agregado puede verse una tabla con el status de la escuela en cada área evaluada y la ganancia promedio en el período considerado. Dicha ganancia es el valor agregado. La única especificación técnica sobre el modo de medición de dicho valor se ubica en el link *Report Card Terms* y dice: “valor agregado mide el progreso de los estudiantes en un grado y tema, lo cual demuestra la influencia que la escuela tiene en el resultado de los alumnos. Este reporte provee información diagnóstica para mejorar las oportunidades educativas para los estudiantes en todos los niveles de desempeño” (Traducción propia).

La tabla referida es la siguiente:

Grades K-8 Value Added -Growth Standard			
(3 years average)*	2009		2009 State
CRT	Status	Mean Gain	Growth Std
Math	F	- 3,3	0
Reading&Language	D	- 0,9	0
Social Studies	B	1-6	0
Science	C	-0,2	0

Nota. Las categorías de status significan: A=ejemplar, B=por encima del promedio, C=promedio, D=debajo del promedio, F=deficiente. (Fuente: “Grade Scale”).

<sup>7</sup> <http://edu.reportcard.state.tn.us/pls/apex/f?p=200:40:2335597636394770>



Más abajo, en la misma página, los resultados se presentan desagregados por año para cada asignatura. A modo de ejemplo se incluyen los resultados de matemática.

Grades K-8 Gain by year, grade, content						
Math						
Grade	2006 - 2007	2007 - 2008	2008 - 2009	3-Year Average	State Growth Standard	State 3 - Year Gain
4th	-3,1	-3,8	-8,9	- 5,3	0	-0,3
5th	- 8	- 0,2	- 2,2	- 3,5	0	-0,1
6th	-2,2	-2,2	-3,9	-1,3	0	0,3
7th					0	0,1
8th					0	-0,4

El programa EVAAS for K-12 aplica el modelo de Tennessee y puede comprarse por internet.<sup>8</sup> En la página en la que se comercializa el programa se promueve su uso entre docentes proactivos para que, “identificando cuáles alumnos están en riesgo (...), puedan asegurar que todos tengan chance de tener éxito” (Traducción propia).

### 3.2. Valor agregado en Inglaterra

En la página del departamento de educación del gobierno de Inglaterra se puede acceder a los reportes de valor agregado de cada centro educativo.<sup>9</sup>

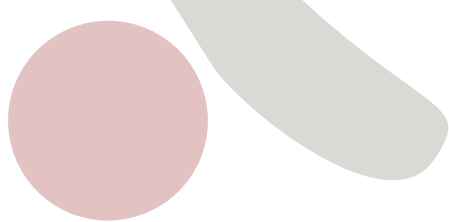
La ventana principal del Departamento de Educación, donde se encuentran las tablas de logro de las escuelas de Inglaterra, permite seleccionar el nivel de la escolaridad para el cual se quiere la información, así como seleccionar una institución o autoridad local en particular.

Al seleccionar una autoridad local aparece una ventana con los resultados de todas las escuelas comprendidas en ella y con varias pestañas que estructuran la información. En la pestaña valor agregado contextualizado se encuentra un ranking de escuelas según esta medida, para el grado seleccionado, donde además se presenta el intervalo de confianza de la medida y la cobertura alcanzada.

<sup>8</sup> <http://www.sas.com/govedu/edu/k12/evaas/index.html>

<sup>9</sup> [http://www.education.gov.uk/performancetables/primary\\_09.shtml](http://www.education.gov.uk/performancetables/primary_09.shtml)





Si bien en la tabla no se indica cómo está calculado el valor agregado contextualizado, en el mismo sitio web se presenta información al respecto, con distinto grado de sofisticación. En términos generales, se dice que esta medida

“toma en cuenta los puntos de partida de cada alumno en los resultados del test (...) y también ajusta por factores que están fuera del control de las escuelas (como el sexo, movilidad y niveles de deprivación) que se ha visto impactan en los resultados de los alumnos” (Department for Children School and Families, 2009:1. Traducción propia).

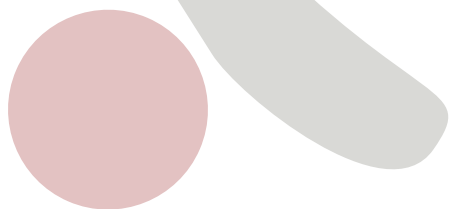
Además, en el link de guía para usuarios y recursos, en la opción documentos, se encuentran tres de ellos que detallan la forma de cálculo: ready reckoner mainstream schools.xls, ready reckoner special schools.xls y technical guide to CVA 2009 model.doc.<sup>10</sup>

A continuación se incluye el ranking de escuelas según su valor agregado contextualizado.

Background	Contextual Value Added	KS2 Results	Year on year Comparisons	Progress measures	Absence
		Contextual Value Added			
		KS1-2 CVA Score	CVA Confidence interval		Coverage
			Lower Limit	Upper Limit	
Local Authority Average					
England (maintained schools only)					
England (all schools)					
The Abbey RC Junior and Infant School - Birmingham		99.3	98.8	99.8	97%
Acocks Green Primary School - Birmingham		100.8	100.3	101.3	97%
Adderley Primary School - Birmingham		100.7	100.2	101.2	69%
Al-Furqan Primary School - Birmingham		99.5	98.9	100.2	100%
Albert Bradbeer Primary - Birmingham		99.8	99.2	100.3	100%
Allens Croft Primary School - Birmingham		100.5	99.8	101.2	100%
Alston Primary School - Birmingham		100.3	99.9	100.7	94%
Anderton Park Primary School - Birmingham		100.0	99.5	100.4	97%
Anglesey Primary School - Birmingham		100.8	100.4	101.2	97%
Arden Primary School - Birmingham		100.8	100.3	101.3	90%
Aston Tower Community Primary School - Birmingham		99.4	98.9	100.0	92%
Audley Primary School - Birmingham		101.0	100.5	101.5	96%
Banners Gate Junior School - Sutton Coldfield		98.9	98.3	99.4	96%
Barford Primary School - Birmingham		99.4	98.9	100.0	92%
Bellfield Junior SU School - Birmingham		100.0	99.5	100.6	94%

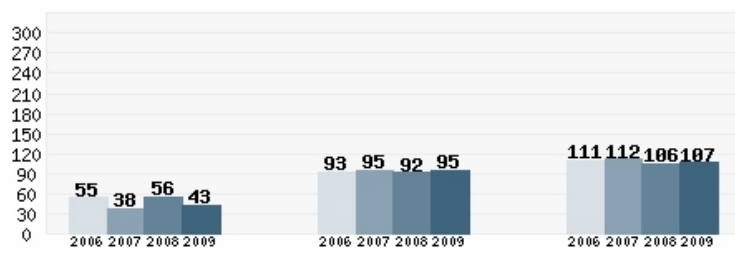
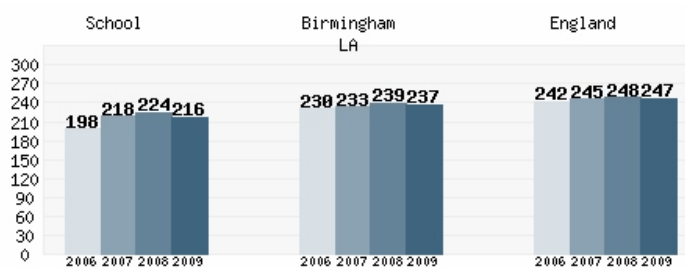
En otra de las pestañas se presentan las medidas de progreso de cada escuela. En el link [http://www.education.gov.uk/performancetables/primary\\_09/documents.shtml](http://www.education.gov.uk/performancetables/primary_09/documents.shtml) se encuentra un documento guía para la lectura de dichas medidas, entre las cuales se incluye: el porcentaje de alumnos que están haciendo el progreso esperado en inglés y en matemática, así como la cobertura para los dos casos.

<sup>10</sup> [http://www.education.gov.uk/performancetables/primary\\_09/documents.shtml](http://www.education.gov.uk/performancetables/primary_09/documents.shtml)



Background	Contextual Value Added	KS2 Results	Year on year Comparisons	Progress measures	Absence
		% making expected progress in English	Coverage		% making expected progress in maths
			% of eligible pupils included in English progress measure	% of eligible pupils included in maths progress measure	
<b>Local Authority Average</b>					
England (maintained schools only)					
England (all schools)					
The Abbey RC Junior and Infant School - Birmingham		88%	95%	66%	95%
Acocks Green Primary School - Birmingham		85%	100%	88%	100%
Adderley Primary School - Birmingham		88%	65%	93%	67%
Al-Furqan Primary School - Birmingham		68%	100%	88%	100%
Albert Bradbeer Primary - Birmingham		65%	100%	81%	96%
Allens Croft Primary School - Birmingham		86%	100%	79%	97%
Alston Primary School - Birmingham		86%	94%	80%	94%
Anderton Park Primary School - Birmingham		88%	94%	87%	93%
Anglesey Primary School - Birmingham		85%	94%	91%	93%
Arden Primary School - Birmingham		80%	90%	77%	88%
Aston Tower Community Primary School - Birmingham		84%	90%	79%	88%
Audley Primary School - Birmingham		94%	92%	86%	94%

Al seleccionar una escuela en particular, además de la información comparativa con los otros centros educativos, se presentan dos gráficos en donde se compara la escuela con las otras de la misma localidad y con el total del país. En el documento *Achievement and Attainment Tables* se explica que “las gráficas de barras muestran la suma de los tres porcentajes de alumnos que alcanzan el nivel 4 y más, y el nivel 5, en inglés, matemática y ciencias, en cada uno de los cuatro años anteriores” (Department for Children School and Families, 2009:3. Traducción propia).



El programa RAISEonline (*Reporting and Analysis for Improvement through School Self-Evaluation*)<sup>11</sup> es una herramienta de análisis de datos para que cada escuela de Inglaterra pueda realizar una autoevaluación. Fue “desarrollado conjuntamente por el gobierno central y la agencia independiente de inspección de las escuelas” (Ray et al. 2009:47. Traducción propia).

### 3.3. Un ejemplo de ranking de docentes y escuelas en Estados Unidos

Una diferencia importante entre Estados Unidos e Inglaterra es que en el primero algunos estados elaboran un ranking de profesores, lo cual no sucede en Inglaterra, donde el valor agregado de cada docente no está disponible a nivel público, pero sí en cada escuela (Ray, 2009:62).

<sup>11</sup> <https://www.raiseonline.org>



Si bien en la tabla no se indica cómo está calculado el valor agregado contextualizado, en el mismo sitio web se presenta información al respecto, con distinto grado de sofisticación. En términos generales, se dice que esta medida

“toma en cuenta los puntos de partida de cada alumno en los resultados del test (...) y también ajusta por factores que están fuera del control de las escuelas (como el sexo, movilidad y niveles de deprivación) que se ha visto impactan en los resultados de los alumnos” (Department for Children School and Families, 2009:1. Traducción propia).

Además, en el link de guía para usuarios y recursos, en la opción documentos, se encuentran tres de ellos que detallan la forma de cálculo: ready reckoner mainstream schools.xls, ready reckoner special schools.xls y technical guide to CVA 2009 model.doc.<sup>10</sup>

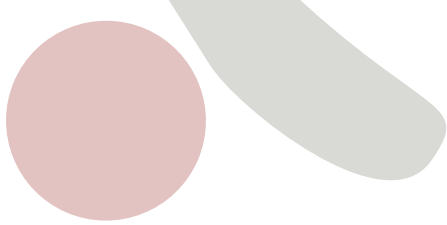
A continuación se incluye el ranking de escuelas según su valor agregado contextualizado.

Background	Contextual Value Added	KS2 Results	Year on year Comparisons	Progress measures	Absence
		Contextual Value Added			
		KS1-2 CVA Score	CVA Confidence interval		Coverage
			Lower Limit	Upper Limit	
Local Authority Average					
England (maintained schools only)					
England (all schools)					
The Abbey RC Junior and Infant School - Birmingham		99.3	98.8	99.8	97%
Acocks Green Primary School - Birmingham		100.8	100.3	101.3	97%
Adderley Primary School - Birmingham		100.7	100.2	101.2	69%
Al-Furqan Primary School - Birmingham		99.5	98.9	100.2	100%
Albert Bradbeer Primary - Birmingham		99.8	99.2	100.3	100%
Allens Croft Primary School - Birmingham		100.5	99.8	101.2	100%
Alston Primary School - Birmingham		100.3	99.9	100.7	94%
Anderton Park Primary School - Birmingham		100.0	99.5	100.4	97%
Anglesey Primary School - Birmingham		100.8	100.4	101.2	97%
Arden Primary School - Birmingham		100.8	100.3	101.3	90%
Aston Tower Community Primary School - Birmingham		99.4	98.9	100.0	92%
Audley Primary School - Birmingham		101.0	100.5	101.5	96%
Banners Gate Junior School - Sutton Coldfield		98.9	98.3	99.4	96%
Barford Primary School - Birmingham		99.4	98.9	100.0	92%
Bellfield Junior SU School - Birmingham		100.0	99.5	100.6	94%

En otra de las pestañas se presentan las medidas de progreso de cada escuela. En el link [http://www.education.gov.uk/performance/tables/primary\\_09/documents.shtml](http://www.education.gov.uk/performance/tables/primary_09/documents.shtml) se encuentra un documento guía para la lectura de dichas medidas, entre las cuales se incluye: el porcentaje de alumnos que están haciendo el progreso esperado en inglés y en matemática, así como la cobertura para los dos casos.

Background	Contextual Value Added	KS2 Results	Year on year Comparisons	Progress measures	Absence
		% making expected progress in English	Coverage		% making expected progress in maths
			% of eligible pupils included in English progress measure	% of eligible pupils included in maths progress measure	
Local Authority Average					
England (maintained schools only)					
England (all schools)					
The Abbey RC Junior and Infant School - Birmingham		88%	95%	66%	95%
Acocks Green Primary School - Birmingham		85%	100%	88%	100%
Adderley Primary School - Birmingham		88%	65%	93%	67%
Al-Furqan Primary School - Birmingham		68%	100%	88%	100%
Albert Bradbeer Primary - Birmingham		65%	100%	81%	96%
Allens Croft Primary School - Birmingham		86%	100%	79%	97%
Alston Primary School - Birmingham		86%	94%	80%	94%
Anderton Park Primary School - Birmingham		88%	94%	87%	93%
Anglesey Primary School - Birmingham		85%	94%	91%	93%
Arden Primary School - Birmingham		80%	90%	77%	88%
Aston Tower Community Primary School - Birmingham		84%	90%	79%	88%
Audley Primary School - Birmingham		94%	92%	86%	94%





El uso del valor agregado para rankear a los docentes es un aspecto muy controversial y, como se verá más adelante, es considerado como una de las mayores debilidades que tienen estos modelos. Sin embargo, recientemente en Los Ángeles se publicó un ranking de docentes<sup>12</sup> que, llamativamente, no presenta ninguna contextualización de los resultados. En la página se puede acceder a la calificación de valor agregado para cada docente (en matemática e inglés y en un promedio de ambas), así como al nivel de cada centro educativo. También se incluye una lista de los 100 mejores docentes y los 100 mejores centros<sup>13</sup>. En los resultados de los centros se puede ver, además de la ubicación en la escala de *valor agregado*, el porcentaje de alumnos que alcanzan los estándares estatales.

En la sección de preguntas frecuentes se encuentra un link al reporte técnico de los análisis realizados para elaborar el ranking<sup>14</sup> y también a recursos sobre valor agregado<sup>15</sup>.

Entre los resultados presentados en dicho reporte, se señala que

“la mayoría de los factores de contexto de los estudiantes no están relacionados con la efectividad docente (...) Tres factores tienen efecto significativo sobre los puntajes de valor agregado, pero la magnitud de esos efectos es pequeña (...) Estos efectos pequeños sugieren que la medida de valor agregado está haciendo un buen trabajo controlando por la diversidad de alumnos asignados a docentes individuales” (Buddin, s/f: 15. Traducción propia).

La imagen que sigue muestra la opción para buscar docentes o centros educativos, así como la lista de los 100 mejores en cada caso.

**Find a teacher...**

**Or, find a school**

### Información para cada docente

Al indicar un docente en particular, se accede a una página en la que se aclara que el puntaje del valor agregado de cada docente está basado en el progreso de sus alumnos en matemática e inglés, y que se tomó toda la información disponible entre 2002-2003 a 2008-2009, así como que dicha medida “captura solo un aspecto del trabajo docente”.<sup>16</sup>

Como puede observarse, lo más destacado de la página son las tres barras horizontales en donde se ubica con un rombo al docente, tanto en matemática como en inglés, y en un promedio de ambas. Según la ubicación del rombo es posible clasificar al docente en cinco categorías desde el menos efectivo al más efectivo.

<sup>12</sup> Aproximadamente 6000 docentes en 470 escuelas.

<sup>13</sup> <http://projects.latimes.com/value-added/>

<sup>14</sup> <http://www.latimes.com/media/acrobat/2010-08/55538493.pdf>

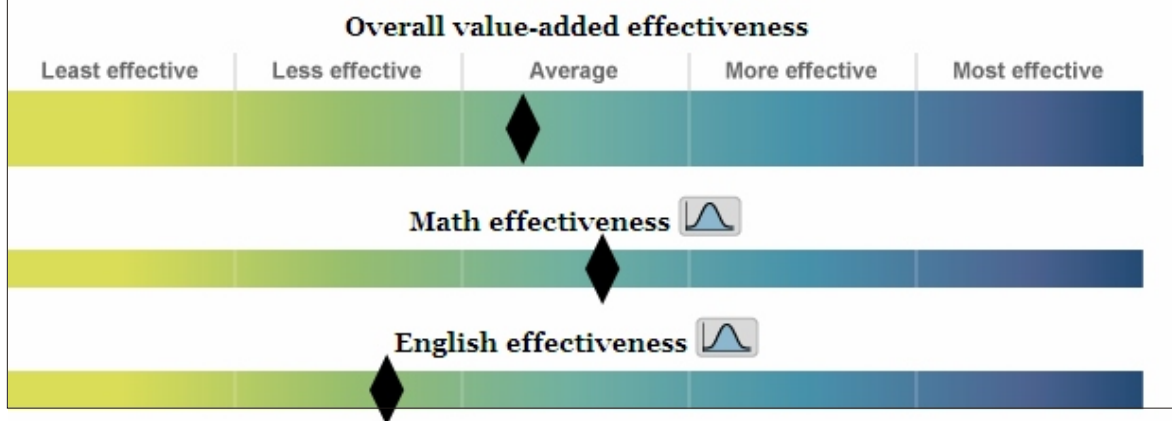
<sup>15</sup> <http://projects.latimes.com/value-added/resources>

<sup>16</sup> <http://projects.latimes.com/value-added/teacher/>

## Edith Aileen Close Vaziri

A 3rd grade teacher at [Elysian Heights Elementary](#) in 2009

These graphs show a teacher's "value-added" rating based on his or her students' progress on the California Standards Tests in math and English. The Times' analysis used all valid student scores available for this teacher from the 2002-03 through 2008-09 academic years. The value-added scores reflect a teacher's effectiveness at raising standardized test scores and, as such, capture only one aspect of a teacher's work.



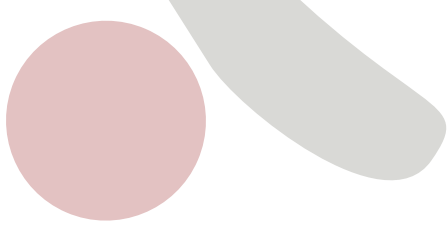
Luego del gráfico se presenta una interpretación de este aclarando que el docente ranqueó (en este caso): (a) promedio en general; (b) promedio en matemática, -lo cual implica que- los estudiantes de los docentes en esta categoría, en promedio, no ganaron o perdieron significativamente en los estándares del estado comparados con otros estudiantes en el mismo grado; y (c) menos efectivo que el promedio en inglés, -lo cual implica que los alumnos de los docentes en esta categoría, en promedio, perdieron alrededor de tres puntos percentiles en los estándares del estado comparados con otros alumnos en su grado.

Junto con esta información se encuentra un link de *historia docente*, que remite a la información de valor agregado de la escuela (que se muestra más adelante). Al final se presenta una tabla que, si bien tiene por objetivo mostrar en qué grado enseñó cada docente desde el año 2003 a 2009, los ordena según su valor agregado, lo cual en sí constituye un *ranking docente* al interior de cada escuela.

Teacher	School years / Grade taught						
	2009	2008	2007	2006	2005	2004	2003
<b>Most effective value-added teachers</b>							
Mc Lean, Gerald	3rd	3rd	3rd	3rd	3rd	3rd	3rd
Saballos, Alina	4/5	2nd	3rd	3rd	3rd	3rd	3rd
<b>More effective than average value-added teachers</b>							
Ermatinger, Susan	-	4/5	5th	5th	4th	4th	3rd
Hawkins, Amy				4th	4th	4th	4th
Ortiz, Cherry	4th	5th	-	3rd	3rd	3rd	3rd
<b>Average value-added teachers</b>							
Close Vaziri, Edith	3rd	3rd	4/5	5th	5th	5th	5th
Martinez, Ruben					5th		
<b>Less effective than average value-added teachers</b>							
Szabo, Steve	-	4th	4th	4th	4th	4th	4th

Cabe preguntarse si este tipo de presentación de resultados contribuye o no a la mejora escolar, ya que es probable que presentar la información de esta manera juegue en contra del clima escolar, aún en centros educativos en los que haya un buen clima de trabajo liderado por la dirección y en donde se observen altos niveles de confianza interpersonal entre los docentes.

Una vez más, este ejemplo ilustra cómo el uso que se hace de la información puede no ser el más adecuado para los objetivos perseguidos, por mejores que sean las medidas estadísticas construidas.

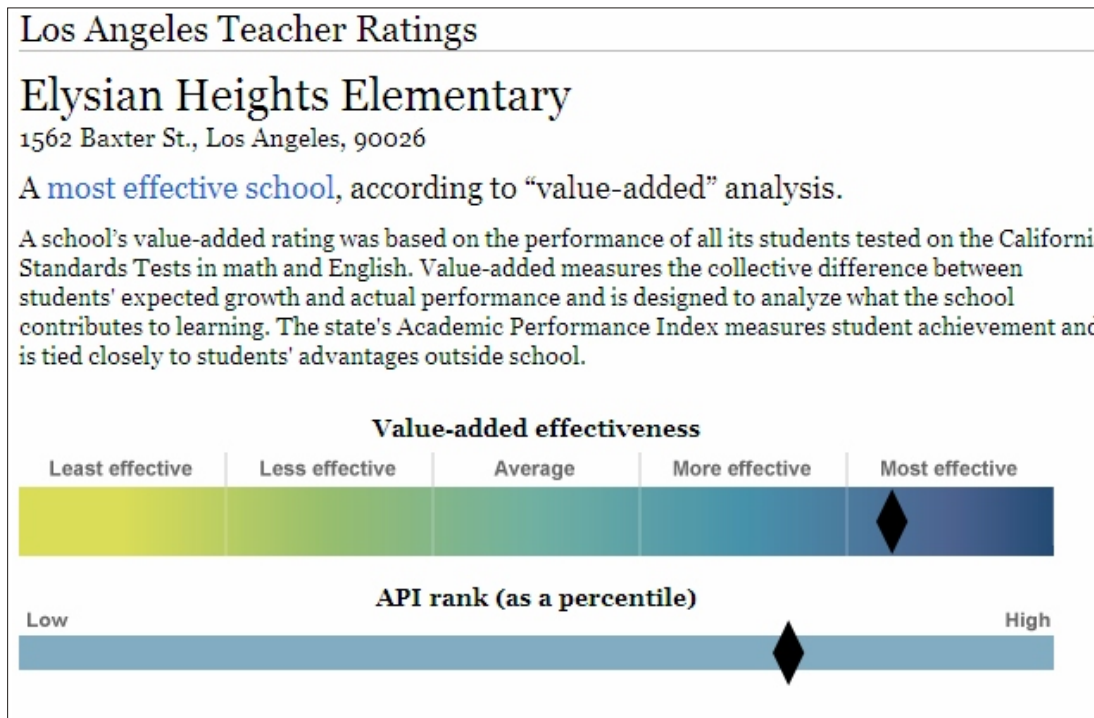


**Información para cada escuela**

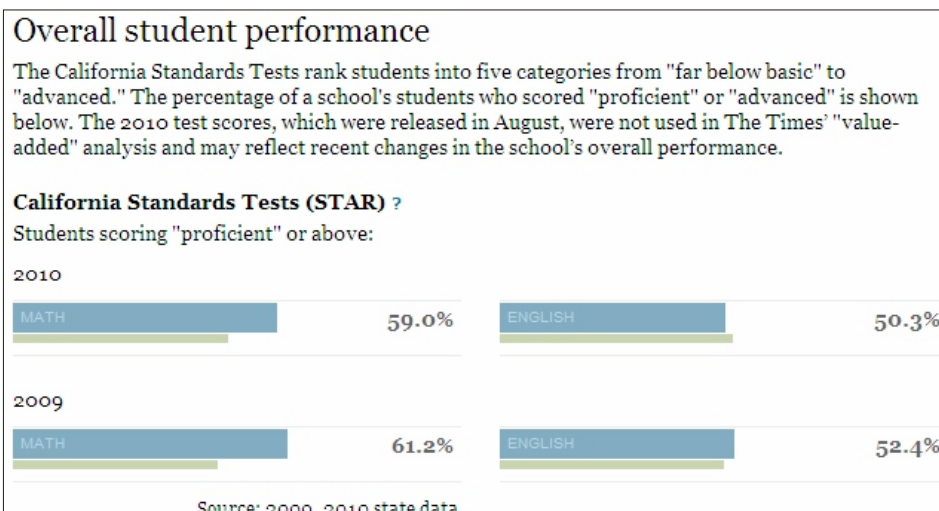
Para cada escuela, la información de valor agregado disponible es similar a la que se presenta para cada docente, pero sin diferenciar entre áreas. Además, se presenta el Índice de Desempeño Académico (API), del cual se afirma que mide el resultado de los estudiantes y está estrechamente relacionado con las ventajas de los alumnos fuera de la escuela.

En la primera barra es posible ubicar a cada escuela según su valor agregado. En la segunda barra, la ubicación de cada escuela surge del índice API<sup>17</sup>, el cual, mide el desempeño y no el valor agregado. La comparación entre ambas barras puede iluminar sobre la mejora escolar (valor agregado) y a la vez el nivel de desempeño de los alumnos. Teóricamente sería posible que escuelas “muy efectivas” tengan un API *relativamente bajo*: se trataría de escuelas con desempeños más bajos que el resto (API), pero que han logrado mejorar respecto a sí mismas en relación a las mediciones anteriores (valor agregado). En el otro extremo habría escuelas “muy efectivas” (que han mejorado en relación a sí mismas) y que a su vez, sus alumnos se ubican entre quienes obtienen los puntajes más altos en las pruebas (API).

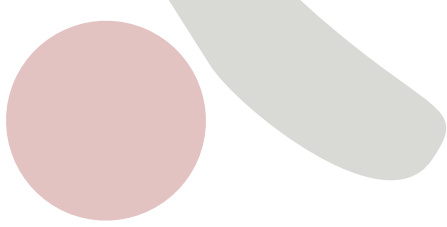
Ambos aspectos brindan información diferente y complementaria: cuánto mejora la escuela en relación a sí misma y cómo se ubica en relación a las demás según el desempeño de sus alumnos.



Más abajo se incluye el porcentaje de alumnos de la escuela que alcanza o supera, independientemente de su valor agregado, los estándares establecidos por el estado en inglés y matemática, para los años 2010 y 2009. La barra celeste representa el porcentaje correspondiente a la escuela y la gris el que corresponde al estado.



<sup>17</sup> Academic Performance Index (API).



## Opinión de los docentes

A lo largo de estas páginas existen espacios disponibles para que los docentes escriban su opinión sobre preguntas tales como: ¿Ud. piensa que el valor agregado puede ser usado para ranquear docentes? ¿Por qué sí o por qué no?, ¿Los puntajes en esta base reflejan su experiencia o la de sus alumnos en esta escuela? y ¿Ud. cree que esto es una herramienta útil para los padres?

A continuación se transcriben algunos comentarios de docentes. Se seleccionó un docente muy poco efectivo, otro poco efectivo, uno promedio y otro muy efectivo. Los lectores interesados podrán encontrar muchos más en la página de Los Ángeles Times.

Tracy Lynn Evans<sup>18</sup>:

“Evaluar a los profesores de esta manera solo lleva a los docentes a enseñar para el test (...). Mis alumnos tienen excelentes puntajes competentes y avanzados. Incluso algunos tienen puntuaciones perfectas de 600 puntos. Entonces, si esos alumnos mantienen sus puntuaciones perfectas, eso me hace una profesora promedio porque ellos no mejoraron. Pero si ellos pierden y puntúan en 580 o 560, ¡yo voy a ser considerada menos efectiva! (...) Me temo que este artículo solo desalentará a los maestros que ponen su corazón y alma en su trabajo mientras se enfrentan a clases que aumentan su tamaño, desarrollo profesional limitado y disminución de recursos” (Traducción propia).

Por otra parte, más allá de la opinión de la docente, interesa destacar que ella es ranqueada como menos efectiva que el promedio en una escuela (Wilbur Avenue Elementary) que recibe la calificación entre las menos efectivas, pero en donde entre un 75% y 80% de los alumnos alcanzan o superan el estándar del estado en las pruebas de matemática e inglés durante los años 2009 y 2010.

Maria Laura Pérez<sup>19</sup>:

“Es injusto para los docentes ser juzgados según el puntaje de sus alumnos. En primer lugar, los alumnos no son asignados aleatoriamente a las aulas (...) Los alumnos son ubicados en las clases según su nivel de inglés (...) Muchos de mis alumnos tienen 'lagunas' en su aprendizaje y no tienen el vocabulario académico necesario para desempeñarse bien en las pruebas. Sin embargo, el distrito espera de nosotros que sigamos los planes al mismo ritmo a pesar de que muchos de mis alumnos aprenden de manera diferente. Adicionalmente, mientras los estudiantes también aprenden unos de otros, ellos necesitan pares que puedan ser 'modelos'. Cuando los alumnos son agrupados en base a su nivel en inglés, los alumnos en los niveles más bajos carecen de buenos modelos” (Traducción propia).

Ashley Collett Tanger:

“(...) Yo soy evaluada más a menudo que cualquier otro trabajo en la tierra. De acuerdo a su investigación, los únicos datos que interesan para mi evaluación son los cuatro días en que mis estudiantes son evaluados, cuando en realidad yo veo cada día como una evaluación y por supuesto espero encontrar en los ojos de mis estudiantes que soy algo más que 'promedio’” (Traducción propia).

Debra J. Packer<sup>20</sup>:

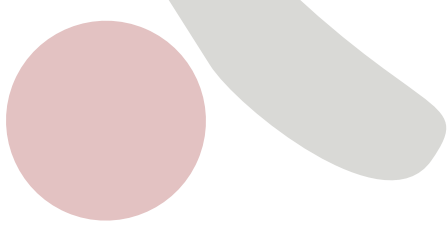
“Como fui clasificada 'muy efectiva' en el reporte del ranking de valor agregado de docentes del LA Times, siento que puedo responder al proyecto sin ser acusada de una mala actitud. Pienso que la mejor manera de aproximarse al absurdo del estudio de LA Times es mostrar una analogía sobre cómo las 'estadísticas' pueden ser malinterpretadas... y consecuentemente causar daños irreversibles debido a la forma sensacionalista en la que se presentaron. Unos pocos años después de que nuestra escuela reabrió, una madre con un muy buen nivel educativo planteó la preocupación de que nosotros no estábamos haciendo lo suficiente para asegurar que nuestros alumnos *superdotados* (*gifted children*) permanecieran en la escuela. Ella habló con otros padres y finalmente fue necesario tratar el tema en la reunión mensual de la escuela (...) Los padres comenzaron la conversación agitando un papel en el aire y declarando: '*nuestro reporte de setiembre muestra que quince niños superdotados han dejado la escuela!*'. Con la cara más seria que pude poner dadas las circunstancias, dije: 'siento mucho eso pero esos niños se graduaron de quinto grado en junio y ese es el motivo por el cual se fueron'. A pesar de que las estadísticas fueron explicadas, hasta hoy esos padres aún cuestionan nuestros programas basados en aquellos cuestionamientos iniciales. Mi punto aquí es que aún gente con un alto nivel educativo frecuentemente puede dejarse llevar por las 'estadísticas'. Las estadísticas dan números y tendencias, pero no incluyen el elemento humano. LA Times ha conseguido desmoralizar a los docentes destacados poniéndoles una etiqueta y como una advertencia en la sección de preguntas frecuentes diciendo que sus hallazgos deben ser solo parcialmente considerados en una evaluación docente general. ¿Cuánta gente se detiene a leer eso, e incluso si lo hacen, su juicio ha sido nublado con 'bueno, si ellos no son buenos en inglés y matemática, ¿cómo pueden ser buenos en algo más?' Los niños no son estadísticas. Los niños son nuestra vida (...). Ellos necesitan ser inspirados para ser grandes científicos y pensadores. ¡Pensar en algunos de nuestros grandes hombres y mujeres que tuvieron problemas de aprendizaje o que pensaban mucho más allá de la 'simplicidad' del inglés y la matemática! Sus profesores probablemente no hubieran pasado el test del valor agregado, pero ellos inspiraron a esa gente y la animaron a seguir sus ideas. Piense en ello, LA Times, antes de ir manipulando nuestro futuro nuevamente” (Traducción propia).

<sup>18</sup> <http://projects.latimes.com/value-added/teacher/tracey-lynn-evans/>

<sup>19</sup> <http://projects.latimes.com/value-added/teacher/maria-laura-perez/>

<sup>20</sup> <http://projects.latimes.com/value-added/teacher/debra-j-packer/>





## 4. Limitaciones de los modelos de valor agregado

Varios autores han señalado limitaciones del valor agregado para identificar efectos, especialmente de docentes. Rivkin considera que

“La miríada de factores que influyen el crecimiento cognitivo, la selección o asignación deliberada de las familias y los docentes en las escuelas y aulas de clase, y las imperfecciones de los test como medidas de conocimiento complican los esfuerzos para estimar el efecto docente” (2007:5. Traducción propia).

En particular se han señalado las siguientes limitantes:

- Como los alumnos no fueron asignados aleatoriamente a cada grupo ni a las escuelas, no parece adecuado identificar el efecto docente con cada uno de ellos en particular. Estos efectos se ven afectados por las características de los alumnos. ¿Cuál es la causa y cuál el resultado? *¿El efecto docente causa el aprendizaje de los alumnos o las características de los alumnos condicionan el efecto docente?* (Wiley, s/f: 11. Traducción propia).
- “Aún si la situación ideal asignación aleatoria de docentes y alumnos- fuera posible, el pequeño número de estudiantes con los que trabaja un docente cada año puede influenciar los efectos estimados de los docentes” (Wiley, s/f:11. Traducción propia).
- “Los cambios anuales en la composición de una clase pueden producir volatilidad año a año, lo cual hace atípico el efecto docente estimado en un año particular” (Wiley, s/f:11. Traducción propia).
- Un aspecto que cobra mayor importancia es la falta de información (*missing data*). Si bien esto es así en toda investigación, en el caso de los estudios que buscan identificar el efecto de un docente o un aspecto de cierta escuela, se vuelve más problemático, ya que la forma que se utilice para solucionar dicha falta de información incidirá sobre la estimación de dicho efecto (Wiley, s/f: 14. Traducción propia).

En la bibliografía sobre el tema se observa un consenso bastante importante acerca de la limitación del valor agregado para estimar efectos de docentes individuales y para tomar decisiones, por ejemplo, sobre salarios.

Wiley considera que

“el efecto docente o escuela estimado a partir de modelos de valor agregado debe ser tomado solo como eso, una estimación (...) Las estimaciones de valor agregado nunca deben servir como el único indicador de efectividad docente y las decisiones importantes no se deben basar primariamente en ellas” (Wiley, s/f:51. Traducción propia).

Esto mismo lo comparte el creador del modelo de Tennessee cuando dice que

“Las estimaciones con valor agregado de la efectividad docente deben ser una parte de la evaluación formal de los profesores, pero no deben ser la única base sobre la cual los docentes sean evaluados, porque existen muchos otros derechos, dimensiones y responsabilidades que no pueden ser medidos por un proceso como el -de TVAAS” (Sanders, 2000:335. Traducción propia).

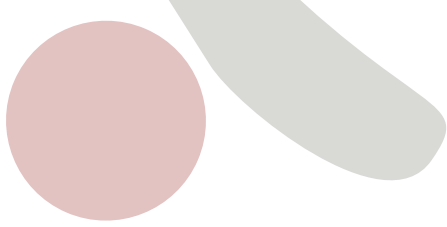
Desde una postura más cautelosa aún, Raudenbush plantea que

“ciertamente las estimaciones de valor agregado, cuando se combinan con otra información, tienen el potencial de estimular discusiones útiles sobre cómo mejorar la práctica. Pero ellas no deben ser tomadas como evidencia directa de los efectos de las prácticas educativas” (2004:128. Traducción propia).

Un estudio de la universidad de Berkeley sobre el uso de los modelos de valor agregado para la evaluación docente muestra que

“los juicios de eficacia de los docentes para un docente en particular pueden variar sustancialmente entre modelos estadísticos, asignaturas dictadas y años. Por otra parte, las características de los estudiantes pueden impactar el ranking de docentes, a veces dramáticamente, aún cuando dichas características fueron controladas estadísticamente en el modelo de valor agregado. Un docente que enseña a alumnos menos aventajados en un curso o año en particular típicamente recibe puntajes más bajos de efectividad que el mismo docente enseñando a alumnos más aventajados en un curso o año distinto” (Newton et al, s/f: 1. Traducción propia).

Si bien esto puede resultar confuso, lo que los autores buscan resaltar es que no alcanza con controlar según las características de los alumnos para que la medida de efectividad (valor agregado) de cada docente sea comparable. A partir de sus análisis encuentran que, a pesar de que se establecen controles estadísticos que toman en cuenta la particularidad de cada alumno, el valor agregado que obtienen los docentes que enseñan a niños con desempeños iniciales más bajos tiende a ser menor que el valor agregado de quienes enseñan a niños con desempeños iniciales mejores. En otras palabras, es más fácil hacer que avancen los alumnos que parten de una situación más ventajosa.



Por último, el National Comprehensive Center for Teacher Quality plantea varios consejos y limitaciones en el uso del valor agregado para calificar docentes (Goe, L., 2008:5-7):

- Es insuficiente para determinar el impacto de determinadas prácticas docentes sobre el progreso de los alumnos y dar lineamientos concretos de mejora.
- Es difícil aislar la contribución de un docente en particular.
- No es claro cuán precisos son los modelos para aislar el efecto del docente actual y separar el efecto de los anteriores.
- La falta de información en las bases de datos y los pequeños tamaños de muestra afectan las estimaciones de manera distinta para cada docente.
- Las diferencias en los modelos de valor agregado usados por diferentes estados o distritos influye en el valor agregado de cada docente. Dicha medida no es comparable entre distritos. Es posible que un docente promedio de un distrito haga una mayor contribución al aprendizaje de sus alumnos que un docente muy efectivo de otro.
- No está explícito cuáles son los parámetros de buena enseñanza. Es posible que un docente con un buen valor agregado haya sido bueno no en enseñar el currículum sino en preparar a los alumnos para responder pruebas estandarizadas. También es posible que quien haya enseñado un currículum alineado con la prueba obtenga mejor valor agregado que otro que enseñó un currículum más rico e incluye materiales que no son evaluados por la prueba.

---

## 5. Posibilidades de uso

Los modelos de valor agregado constituyen un avance con relación al uso de resultados directos de los estudiantes (véase al respecto el Boletín N° 16 sobre los semáforos de Chile). El valor agregado es una forma más apropiada de medir el efecto de escuelas y maestros, en la medida en que tiene en cuenta los puntos de partida de los alumnos y los datos del contexto social.

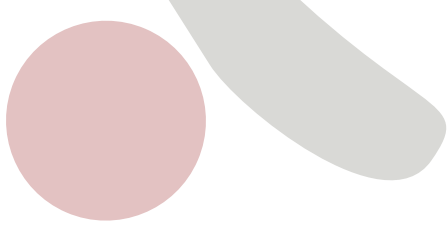
Como toda medida estadística, tiene limitaciones, por lo cual no debe utilizarse en forma aislada, sino en conjunto con otras fuentes de evidencia empírica, sobre todo cuando su uso tiene consecuencias públicas y directas para docentes y escuelas.

En este sentido, la experiencia de autoevaluación de las escuelas y docentes en Inglaterra parece prometedora. Al mismo tiempo, los ranking públicos de docentes elaborados por Los Angeles Times es un ejemplo de mala práctica, dado que emplea el modelo con fines sensacionalistas que no aportan nada a la mejora del trabajo docente.

Un dato alentador es que, recientemente, la Fundación de Bill Gates ha comenzado a financiar la filmación de clases de docentes, de modo de producir información que pueda ser vinculada con las medidas de valor agregado, con el fin de comprender mejor qué aspectos de la enseñanza contribuyen a mejorar el aprendizaje, así como a establecer valoraciones más apropiadas del trabajo de los docentes.

En el siguiente vínculo se puede acceder a información sobre esta nueva iniciativa, tanto en inglés como en castellano:

[http://mt.educarchile.cl/MT/jjbrunner/archives/2010/12/evaluar\\_el\\_dese.html](http://mt.educarchile.cl/MT/jjbrunner/archives/2010/12/evaluar_el_dese.html)



## Referencias

- Brown Center on Education Policy at Brookings (2010) “Evaluating Teachers: The Important Role of Value-Added”.  
[Http://www.brookings.edu/reports/2010/1117\\_evaluating\\_teachers.aspx](http://www.brookings.edu/reports/2010/1117_evaluating_teachers.aspx)
- Buddin, R. (s/f) “How effective are Los Angeles elementary teachers and schools?” Los Angeles Times.  
[Http://www.latimes.com/media/acrobat/2010-08/55538493.pdf](http://www.latimes.com/media/acrobat/2010-08/55538493.pdf)
- Department for Children, School and Families (2009) “Primary School (Key Stage 2) Achievement and attainment tables, 2009 Birmingham330”  
[Http://www.education.gov.uk/performancetables/primary\\_09/pdf\\_09/330.pdf](http://www.education.gov.uk/performancetables/primary_09/pdf_09/330.pdf)
- Goe, L. (2008) “Key issue: Using value added models to identify and support highly effective teachers” National Comprehensive Center for Teacher Quality. Washington DC.
- Martínez Arias, R. (2009) “Usos, aplicaciones y problemas de los modelos de valor añadido en educación” En Revista de Educación, No.348. Ministerio de Educación, Madrid
- Martínez Arias, R., Gaviria Soto, J. L. y Castro Morera, M. (2009) “Concepto y evolución de los modelos de valor añadido en educación” En Revista de Educación, No.348. Ministerio de Educación, Madrid
- McCaffrey, D.F., Lockwood, J.R., Koretz, D.M. y Hamilton, L.S. (2003) “Evaluating value-added models for teacher accountability”. RAND Education.
- PREAL/GTEE (2010) “La evolución del sistema de inspección de escuelas en el Reino Unido”. Boletín N° 18.
- Raundenbush, S. W. (2004) “What are value-added models estimating and what this imply for statistical practice?” Journal of education and behavioral statistics. Vol 29. Nro. 1.
- Ray, A., Evans, H. y McCormack, T. “El uso de los modelos nacionales de valor añadido para la mejora de las escuelas británicas” En Revista de Educación, No.348. Ministerio de Educación, Madrid
- Rivkin, S. (2007) “Value-Added Analysis and Education Policy” National Center for Analysis of Longitudinal Data in Education Research. The Urban Institute
- Sanders, W. (2000) “Value-Added Assessment from Student Achievement Data: Opportunities and Hurdles” Journal of Personnel Evaluation in Education.
- Saunders, L. (2000) “Understanding schools' use of 'value added' data: the psychology and sociology of numbers” Research Papers in Education 15(3)
- Wiley E. (s/f) “A Practitioner's Guide to Value Added Assessment”. University of Colorado at Boulder.  
[http://nepc.colorado.edu/files/Wiley\\_APractitionersGuide.pdf](http://nepc.colorado.edu/files/Wiley_APractitionersGuide.pdf). Consulta 15/10/2010
- Wright, P., Sanders, W. y Rivers, J. (2006) “Measurement of Academic Growth of Individual Students toward Variable and Meaningful Academic Standards” En Lissitz, R. (Ed.). Longitudinal and Value Added Models of Student Performance. Maple Grove, Minnesota: JAM Press
- Newton, X., Darling-Hammond, L., Haertel, E. y Thomas, E. (s/f) “Value-Added Modeling of Teacher Effectiveness: An Exploration of Stability across Models and Contexts”. Graduate School of Education. Berkeley.  
[Http://gse.berkeley.edu/admin/events/docs/epaa.pdf](http://gse.berkeley.edu/admin/events/docs/epaa.pdf)