

USO DEL VALOR AÑADIDO PARA LA EVALUACIÓN DE SISTEMAS EDUCATIVOS Y PARA LA EVALUACIÓN DE DOCENTES

Dra. María Castro Morera
Universidad Complutense de Madrid
Baja California, 2015



Foro: “Construyendo un sistema educativo de calidad”

La idea del Valor
Añadido (VA)

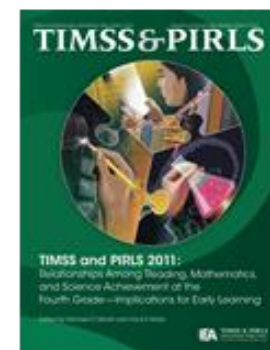
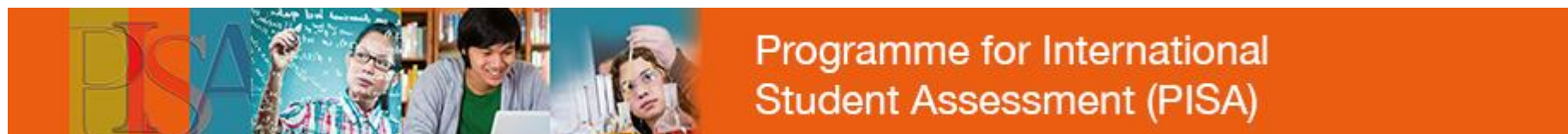
VA y docentes

El impacto en la
eficacia docente

El reto de evaluar
profesores

La idea del valor añadido

Evaluación de sistemas educativos



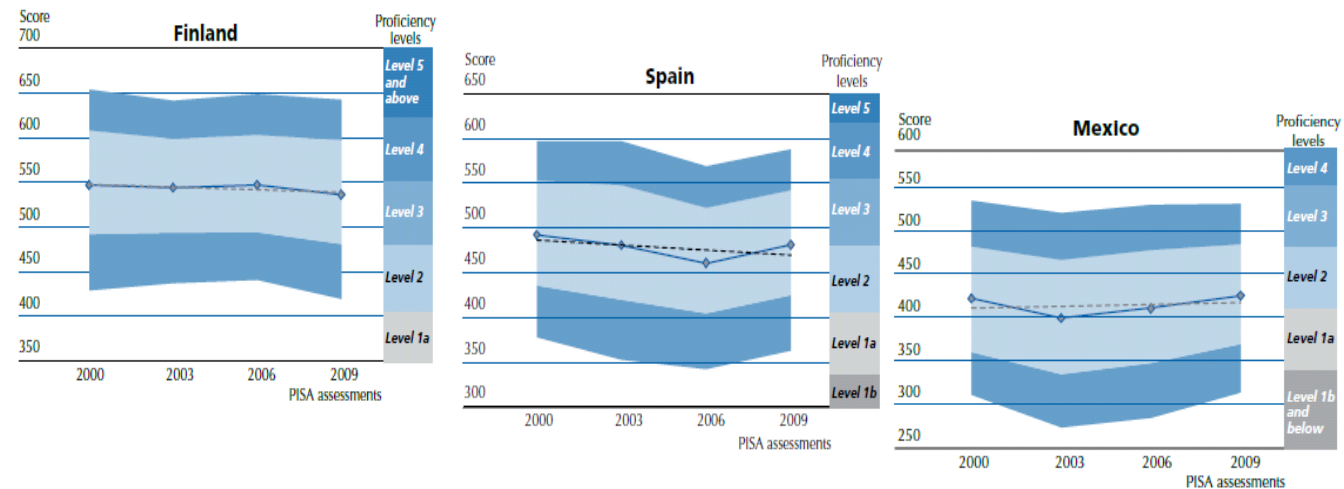
*Instituto Nacional de Estudos e Pesquisas
Educacionais Anísio Teixeira (INEP)*

¿Qué y cuánto saben?
**Evaluaciones de
Status**

Evaluación de sistemas educativos



Las evaluaciones de estatus describen el rendimiento académico de un estudiante o de un grupo de estudiantes en un momento temporal determinado.



Fuente: OEDC (2010). *PISA 2009 results: learning trends. Changes in student performance since 2000. Volumen V.* París: OEDC

¿Qué y cuánto saben?

**Evaluaciones
de Status**

En las evaluaciones de estatus...

Cuando la evaluación de una escuela depende de la comparación del estatus de la cohorte de año anterior con la de este año, entonces:

- Esta comparación confunde la mejora de la escuela con las diferencias entre cohortes.
- Lo que es muy problemático si las escuelas son pequeñas o si tienen alta movilidad de los estudiantes.

¿Qué más evaluar?

¿Qué y cuánto saben?

El **crecimiento** describe el rendimiento académico de un estudiante o de un grupo de estudiantes en dos o más puntos temporales.

El **VA** quiere conocer qué causa el crecimiento de los aprendizajes.

¿Es su evolución adecuada?

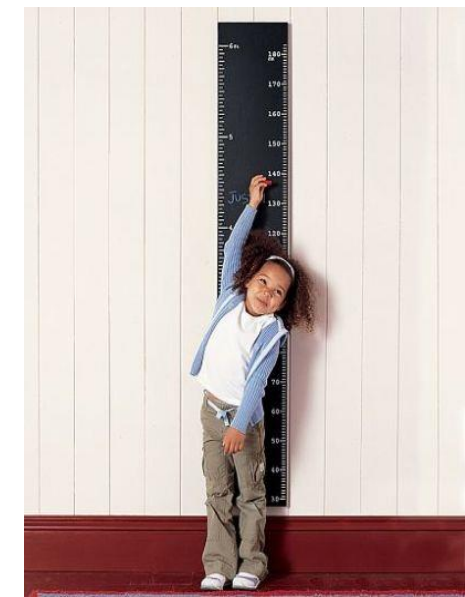
Nivel actual

Tasa de crecimiento

La verdadera
noción de
aprendizaje que
implica
crecimiento y
cambio

Valor añadido

**Crecimiento hacia
un objetivo**



Tasa de crecimiento

¿Saben más o menos que en el momento anterior?

2014

A

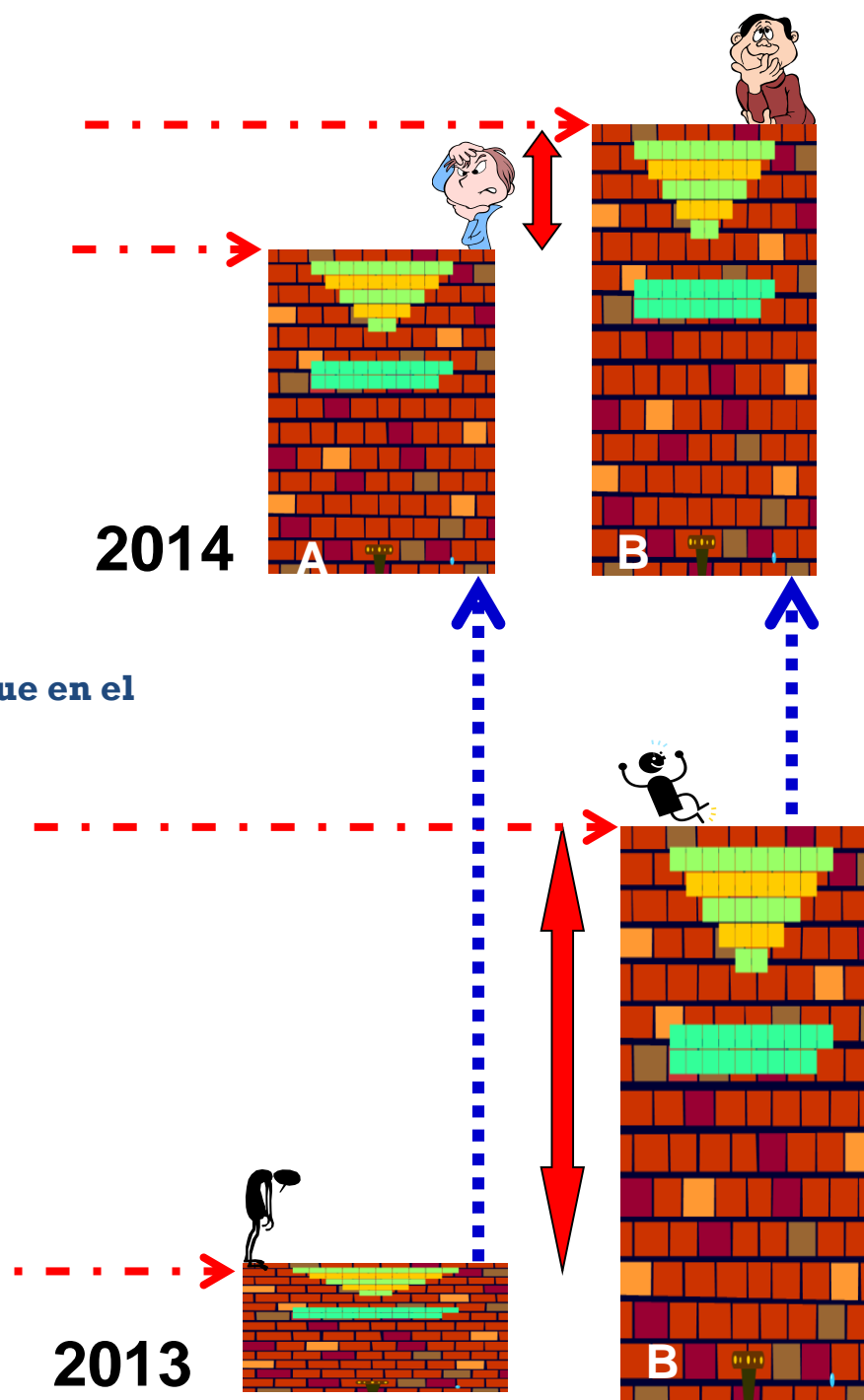
B

Valor añadido

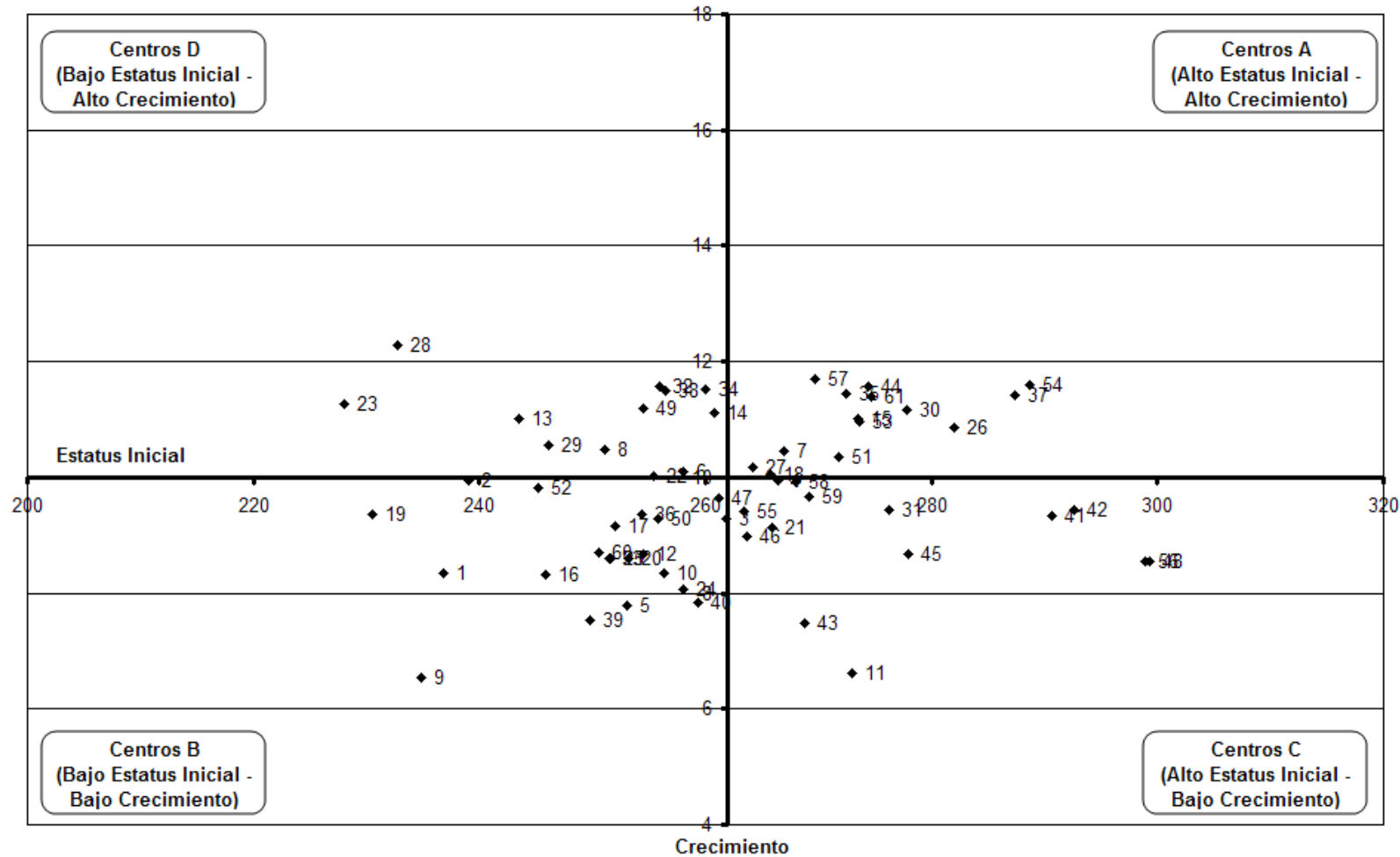
¿Cuál es la aportación de la escuela y del sistema educativo?

2013

B



Estudio del crecimiento en el sistema educativo



Fuente: Castro, M., Ruiz, C. y López, E. (2009). Forma básica del crecimiento en los modelos de valor añadido: Vías para la supresión del efecto de regresión y funciones de crecimiento no lineales. *Revista de Educación*, 348, 116.

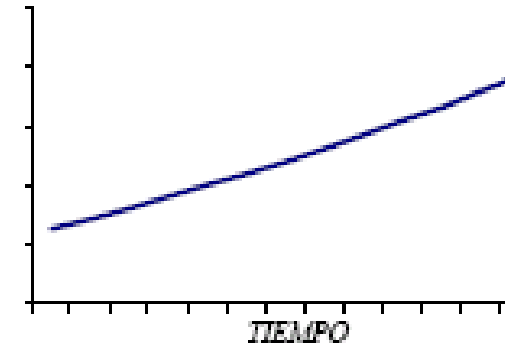
Para la evaluación de escuelas en el marco de los modelos de valor añadido y de crecimiento, es necesario estudiar las diferencias **inter**individuales en los cambios **intra**individuales para establecer regularidades en el desarrollo como función de las prácticas escolares.

(Stevens y Zvoch, 2006)

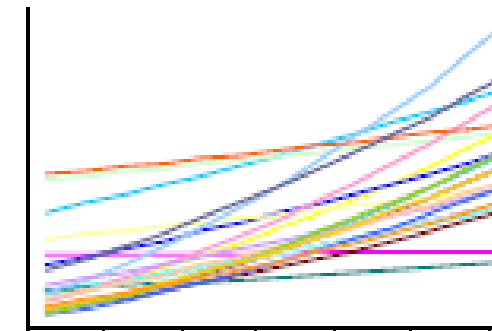


Estudio del crecimiento en el sistema educativo

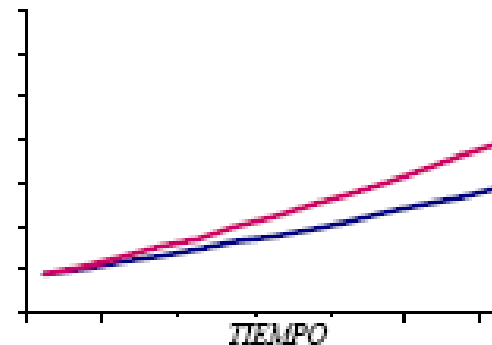
¿Cambia el rendimiento (la conducta) a lo largo del tiempo? ¿Cómo es el cambio intraindividual?
(Nivel 1: Medidas de un sujeto a través del tiempo)



¿Existen diferencia interindividuales en ese cambio? (Nivel 2: Alumnos a través del tiempo)



¿Qué variables pueden explicar esas diferencias?
(Nivel 3: Características de escolares)



En las evaluaciones de crecimiento...

Entonces es cuando se introducen los indicadores de ganancia (de crecimiento) que:

- amplían la base empírica de la evaluación.
- están menos relacionadas con las características de los estudiantes que con el nivel de logro del estudiante.
- reducen el impacto de las diferencias entre cohortes.
- y que mejoran (la percepción) de equidad.

Pero no solucionan el problema de la atribución causal sobre el cambio.

En las evaluaciones de valor añadido...

Se introducen, junto con sus tasas de crecimiento y/o sus niveles de rendimiento anterior, las características contextuales de las escuelas y de los estudiantes. Son modelos contextualizados. Y como en el caso anterior...

tampoco no solucionan el problema de la atribución causal sobre el cambio.

En las evaluaciones de valor añadido...

Los modelos de valor añadido ponen el acento en las ganancias de los estudiantes a través del tiempo. Así, las escuelas más eficaces son aquellas cuyos estudiantes ganan más y las menos eficaces los que ganan menos.



Los modelos de valor añadido ofrecen resultados normativos: las escuelas son comparadas con otras. Aquellas escuelas cuyas contribuciones se determinen como significativamente distintas de la media pueden ser distinguidas para un “tratamiento” especial (premios o sanciones).

Habrá que valorar si desde este tipo de evaluaciones se pueden sostener consecuencias de alto impacto de la evaluación (“*high-stake assessment*”).

Uso e interpretaciones de los resultados de las evaluaciones

- Hay que entender que el nivel de logro y la tasa de crecimiento de los estudiantes son **DESCRIPCIONES** de lo que ocurre en la escuela con respecto a un aspecto del aprendizaje de los alumnos.
- El uso de estas **DESCRIPCIONES** en sistemas de rendición de cuentas (*high-stake assessment*) supone asumir implícitamente que son **indicadores precisos** de la eficacia escolar.

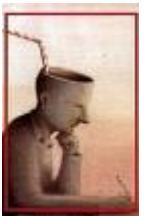
Uso e interpretaciones de los resultados de las evaluaciones

Los indicadores incluidos en las evaluaciones de los sistemas educativos constituyen una interesante **herramienta de “screening”**: pueden usarse para identificar escuelas que **probablemente** realizan grandes contribuciones al aprendizaje de sus alumnos o escuelas que **probablemente** realizan contribuciones más débiles a los aprendizajes.

Los resultados de las evaluaciones (estatus y ganancia) **no** deben utilizarse como base para la toma de decisiones críticas (como la diferenciación de salarios de profesores), aunque pueden considerarse conjuntamente con otras medidas.

Dificultades de los modelos de valor añadido

- Las estimaciones de VA se basan en modelos complejos, que operan con puntuaciones estandarizadas del desempeño de los estudiantes y son de difícil explicación y comprensión para los usuarios del sistema que se muestran críticos con la “opacidad” de los modelos y con la validez de los mismos para tomar decisiones de alto impacto.
- La desventaja conceptual de los modelos de valor añadido está en su operacionalización. El valor añadido parte de una hipótesis (no comprobada completamente) sobre los elementos de la escuela (o características del profesor y de su intervención docente) que producen mejores resultados, y **se ha definido operativamente como un residuo del modelo.**



Valor Añadido y Docentes

La base práctica sobre rendición de cuentas es que es la clave para la mejora de la productividad de las escuelas.

Si las buenas escuelas y los buenos profesores son centrales en el aprendizaje de los estudiantes entonces, ¿la presencia o ausencia de aprendizajes no nos dice algo acerca de la calidad de las escuelas y de los profesores?

(H. Braun, 2007)

Evaluar la calidad de los profesores tomando como referencia el nivel de sus estudiantes al final del año escolar (usando pruebas estandarizadas) es uno de los métodos para evaluar a los profesores, pero favorece a aquellos profesores que cuyos alumnos comienzan con un alto nivel.

El uso de los modelos de valor añadido es una alternativa metodológica para el uso de las puntuaciones estandarizadas en la evaluación de la calidad de la enseñanza.

De modo análogo a lo ya expuesto para el conjunto de los sistemas educativos, se busca aislar la contribución del profesor al crecimiento de los aprendizajes de los estudiantes de otros factores que también influyen (como el rendimiento previo o el nivel socio-económico).



<http://www.carnegieknowledge.org/>

Las propuestas para evaluar profesores basadas en modelos de valor añadido ha generado un intenso debate. Hay tres preocupaciones que subyacen a este debate:

- El sesgo.
- La precisión.
- La relevancia.

Los dos primeros son importantes, de naturaleza técnica. Pero, incluso si las estimaciones de valor añadido fueran “casi perfectas”, su utilidad para evaluar la docencia depende de la relevancia.

Y la relevancia hace referencia a la finalidad de la escolaridad (funcionalidad). En el fondo, ayudar a los estudiantes a desempeñar bien en un test tiene poco valor en si mismo. La pregunta crucial es si las ganancias en los aprendizajes representan aprendizaje y crecimiento en las habilidades y competencias necesarias (y cambiantes) a lo largo de la vida.



Carnegie Foundation for the Advancement of Teaching

WHAT DO WE KNOW ABOUT THE LONG-TERM IMPACTS OF TEACHER VALUE-ADDED?

STEPHEN W. RAUDENBUSH
UNIVERSITY OF CHICAGO

CARNEGIE KNOWLEDGE NETWORK
What We Know Series:
Value-Added Methods and Applications

ATIL

ADVANCING TEACHING - IMPROVING LEARNING



Universidad
Complutense
Madrid

El impacto de la eficacia docente

Magnitud de los efectos iniciales del VA

- La investigación coincide en señalar que asistir a la clase de un profesor que está 1 sd sobre sus iguales en VA está asociado a una ganancia en rendimiento que oscila entre el 10-15% de 1sd en el rendimiento del alumno.



Persistencia de los efectos del VA

- ¿Cuánta diferencia se mantiene tras un año?
- La literatura señala el impacto del VA decae con el tiempo en el rendimiento de las puntuaciones de los test.
- Los más pesimista: 18-25% de la diferencia.
- Los más optimista: 50%
- La mediana está en 24%, lo que significa que $\frac{1}{4}$ del VA inicial permanece después de un año.

Impactos en la edad adulta

- Impactos: asistencia a la universidad, calidad de la universidad (a los 20 años), ingresos (a los 28 años), lugar de residencia y paternidad en la adolescencia.
- La experiencia docente es capital y las habilidades del docente son claves en este conjunto de impactos.

Table 1: Persistence of Value-Added After Initial Year as Fraction of Value-Added During the Initial Year

Study	Sample	Year 1	Year 2	Year 3	Year > 3
Kinsler (2012) ^v	N=689,641 students, grades 3-5, 1998-2005, in North Carolina	.24 (math) .14 (reading)			
Master, Loeb, and Wycoff, 2014	N=700,000 students, grades 3-8, 2005-2014 in New York City	.19 (math) .21 (language arts)			
McCaffrey et al. (2004) ^{vi}	N=678, grades 3-5, large suburban district	.25	.15	--	--
Lockwood et al. ^{vii}	N=10,000, grades 1-5, large urban district	.18 	.15	.14	.12
Kane and Staiger (2008) ^{viii}	97 pairs of teachers, grades 2-5, randomization to students to teachers within pairs	.50			
Jacob, Lefgren, and Sims (2010) ^{ix}	n=18,240, grades 4-15, mid-size Western District	.20			
Rothstein (2010) ^x	n=99,071, grades 3-5, North Carolina statewide	.27 (math) .33 (reading)			
Measurement of Effective Teaching (2012)	1811 teachers randomized within schools to student rosters, grades 4-8 in 6 school districts	.45			
Chetty et al. (2012)	10,992 students randomized to classes within 79 schools in Tennessee				0
Chetty et al. (2013)	2.5 million children grades 3-8 in New York	.50 	.40	.20	.20

Magnitud de los efectos iniciales del VA

- La investigación coincide en señalar que asistir a la clase de un profesor que está 1 sd sobre sus iguales en VA está asociado a una ganancia en rendimiento que oscila entre el 10-15% de 1sd en el rendimiento del alumno.

Persistencia de los efectos del VA

- ¿Cuánta diferencia se mantiene tras un año?
- La literatura señala el impacto del VA decae con el tiempo en el rendimiento de las puntuaciones de los test.
- Los más pesimista: 18-25% de la diferencia.
- Los más optimista: 50%
- La mediana está en 24%, lo que significa que $\frac{1}{4}$ del VA inicial permanece después de un año.

Impactos en la edad adulta

- Impactos: asistencia a la universidad, calidad de la universidad (a los 20 años), ingresos (a los 28 años), lugar de residencia y paternidad en la adolescencia.
- La experiencia docente es capital y las habilidades del docente son claves en este conjunto de impactos.

Table 2: Impacts of Value-Added on Adult Outcomes

	Impact of classroom quality overall (Chetty et al. 2011)	Impact of classroom value-added (Chetty et al. 2011)	Impact of teacher value-added (Chetty et al. 2013)
Initial test scores	8.8 percentiles (.32 sd)		
College Attendance		0.28% above mean of 45.5%	0.82% above mean of 37.22%
College Quality index		0.06 sd	0.02 sd
Earnings	\$1520 (8.8% above mean)	\$1619 (11.1% above mean)	\$350 (1.65% above mean)
Teen parenthood			0.61% below mean of 14.3%
Other outcomes			Increases in neighborhood quality, saving with 401K

El reto de evaluar profesores

REACHStudents

Recognizing Educators Advancing Chicago's Students

- Es una evaluación del desempeño del profesor.
- Crecimiento de los aprendizajes de los alumnos:
 - Performance task
 - Valor añadido
- Práctica profesional (Observations of Teacher Practice) sobre:
 - Planificación y preparación de la clase
 - Entorno del aula
 - Instrucción
 - Responsabilidad profesional

MET

Measurement of Effective Teaching Project

- Es un proyecto de investigación para el desarrollo de medidas de enseñanza efectiva.
- Se emplean múltiples medidas que incluyen estimaciones de VA y también de observación del desempeño del profesor (CLASS).
- Junto con percepción de los alumnos.
- Lo más significativo: se asignan al azar grupos de alumnos a los profesores dentro de cada escuela.

Aún así tenemos los problemas clásicos de la evaluación de sistemas educativos y de profesores:

- La comparación de profesores que trabajan en distintas escuelas confunde la eficacia del profesor con la eficacia de la escuela, con el contexto escolar y con la composición (social) de los alumnos que están escolarizados allí.
- La noción de enseñanza efectiva en estos estudios es completamente relativa. Cuando nos referimos a que algunos profesores son altamente efectivos, sólo estamos diciendo que los profesores dentro de las escuelas son heterogéneos.

Más allá de las consideraciones técnicas, que para nosotros son de capital importancia, es preciso definir cuál es el puerto al que se dirige nuestro sistema educativo, pues sólo con esa definición –que se circunscribe al ámbito social y político- podremos determinar cuan cerca estamos de ese ideal de mejora que es, en definitiva, la función última de la evaluación.

LEY DE CAMPBEL

“The more any quantitative social indicator is used for social decision-making, the more subject it will be to corruption pressures, and the more apt it will be to distort and corrupt the social processes it is intended to monitor”

RECOMENDACIÓN DE BRAUN

“The more poorly an accountability system is designed and/or implemented, the more likely it is to engender resistance and corruption”

(H. Braun, 2007)



¡¡Muchas gracias!!