

Icfes  
Memorias  
del 6º Seminario  
Internacional sobre  
Calidad Educativa  
2015

Elaborado por:  
Jaime Sebastián Lobo Tovar

## Índice

<b>Introducción.....</b>	<b>3</b>
<b>I Antecedentes del seminario .....</b>	<b>4</b>
El seminario anual del Icfes.....	4
Objetivos.....	4
Participantes .....	5
Programa académico, talleres y conferencias .....	5
<b>II Comparación de contenidos de los seminarios 2010 a 2015 .....</b>	<b>8</b>
<b>III Las presentaciones académicas .....</b>	<b>12</b>
Conferencias sobre rendición de cuentas.....	12
Autonomía, toma de decisiones y calidad de la educación, Sussana Loeb.....	12
La alineación de la enseñanza bajo la política “Qué ningún niño se quede atrás” - Morgan Polikoff .....	15
El impacto de las sanciones de la política: “Qué los niños no se queden atrás” en el desempeño de la escuela. Evidencia de una regresión de discontinuidad para Carolina del Norte - Thomas Ahn .....	19
Conferencias sobre valor agregado.....	25
Midiendo el valor agregado en educación superior - Jesse M. Cuhna .....	25
Los primeros siete años en la escuela - Peter Tymms .....	30
Nuevos programas y calidad de la educación superior en Colombia - Juan Esteban Carranza .....	34
Valor agregado en la educación superior - Ernesto San Martín .....	45
Conferencias sobre índices de calidad .....	51
Aprendizaje de monitoreo para todos. Retos y oportunidades de los índices compuestos que sirven a los objetivos de desarrollo sostenible - Silvia Montoya, Unesco.....	51
El IDEB y educación nacional. Retos y perspectivas - Alexandre dos Santos .....	53
Índice sintético de calidad educativa, ISCE - Silvana Godoy y Jorge Duarte .....	56
La experiencia de Medellín en la medición de la calidad educativa. El índice multidimensional de calidad educativa como herramienta para la toma de decisiones - Anderson García .....	58
Índice de la calidad educativa de Bogotá - Luis Alfredo Muñoz.....	61
Sistema de calidad e indicadores de seguimiento para Chile - Carlos Henríquez Calderón .....	63
Sistema de calidad e indicadores de seguimiento para Medellín - Claudia C. Rave.....	67
Las voces de los sujetos. Sistema de monitoreo al Plan Sectorial de Educación - Jorge A. Palacio .....	69
Conferencias Call for Papers.....	70
Brechas de género en el rendimiento escolar a lo largo de la distribución de puntajes. Evidencia de las pruebas Saber 11 - Luz Karime Abadía y Gloria Bernal Nisperuza.....	70

<i>¿Vale la pena prepararse? Efectos de la participación en actividades preparatorias sobre el desempeño en el examen Saber 11 - Silvia Gómez Soler y Paula Herrera.....</i>	<i>71</i>
<i>Predicción automática de la dificultad de preguntas abiertas de respuesta corta con propósitos educativos - George Dueñas.....</i>	<i>71</i>
<i>Valor agregado en lengua extranjera de la educación superior en Colombia - Ramiro Rodríguez Revilla y Diana Carolina López Cuevas.....</i>	<i>72</i>
<i>Factores asociados al desempeño de los estudiantes de psicología en las pruebas Saber Pro - Guillermo Augusto García Méndez.....</i>	<i>73</i>
<i>Etnoeducación y desempeño escolar en la región pacífica colombiana - Fabio Sánchez Torres y Glenda Palacio.....</i>	<i>73</i>
<i>Desigualdad de oportunidades en la educación media y superior en Colombia - Jaime Andrés Sarmiento Espinel y Adriana Carolina Silva Arias.....</i>	<i>74</i>
<i>Efectos de largo plazo del Programa Mamá Canguro sobre resultados educativos y de fuerza laboral - Darío Maldonado.....</i>	<i>74</i>
<i>Educación media y mano de obra industrial en Colombia. Inequidad y pobreza - Jhon James Mora Rodríguez.....</i>	<i>75</i>
<i>¿Qué pasa en el salón de clases con el examen Saber 11? Una etnografía escolar sobre colegios exitosos - Ana María Zamora Moreno, convocatorias Icfes.....</i>	<i>75</i>
<i>Desigualdad de oportunidades en Colombia. Impacto del origen social sobre el desempeño académico y los ingresos de graduados universitarios - Andrea Cuenca.....</i>	<i>76</i>
<i>Convergencia espacial del rendimiento educativo en Colombia (2000-2012) - Osmar Leandro Loaiza Quintero y Guillermo David Hincapié Vélez.....</i>	<i>77</i>
Talleres.....	77
<i>Introducción al análisis estadístico de relaciones causales - Andrés Gutiérrez.....</i>	<i>78</i>
<i>Introducción al uso de modelos estructurales usando MPLUS - Lina Saldarriaga.....</i>	<i>78</i>
<i>Introducción a la evaluación de impacto - Luis Carlos Reyes.....</i>	<i>79</i>
<i>Uso de base datos - Fernanda Zarate, Sandra Roperó y Lorena Wilches.....</i>	<i>79</i>
<b>IV Los participantes y su evaluación sobre el seminario.....</b>	<b>80</b>
<b>V Conclusiones.....</b>	<b>83</b>
Referencias.....	85

## Introducción

Este informe contiene las memorias del 6º Seminario Internacional de Investigación sobre Calidad de la Educación sobre Índices de Calidad, organizado por el Instituto Colombiano para la Evaluación de la Educación (Icfes) y realizado en la ciudad de Bogotá durante la semana del 3 al 6 de noviembre de 2015.

El Seminario Internacional de Investigación sobre la Calidad de la Educación tiene por objetivo dar a conocer investigaciones nacionales e internacionales, así como los estudios en desarrollo por el Icfes sobre la calidad de la educación en Colombia. Este evento es de carácter anual y está dirigido a investigadores, diseñadores de políticas públicas, directivos, docentes y estudiantes de educación superior, a quienes además se les hace la invitación a presentar trabajos de investigación sobre calidad de la educación. A partir de 2015, el Seminario Internacional de Investigación sobre la Calidad de la Educación se enfocará en un tema específico de educación. Para este año, el tema a tratar serán los índices de calidad y reunirá a conferencistas de diferentes nacionalidades que pretenden abarcar el tema desde distintos puntos de vista.

En el primer capítulo del documento se describirán los antecedentes del seminario. El segundo comparará los contenidos de los seminarios anuales 2010 a 2015. El tercer apartado es un resumen ejecutivo de las principales ideas desarrolladas en cada conferencia, tanto en las sesiones principales como en las alternas y en el simposio. En caso de querer profundizar en alguno de los temas presentados, podrá encontrar en las referencias de este documento los artículos relacionados como complemento a la información publicada en la página web <http://www.icfes.gov.co/index.php/component/k2/item/1751?Itemid=594>. En el capítulo cuatro se recogen las conclusiones del seminario.

## I

## Antecedentes del seminario

### El seminario anual del Icfes

La misión del Instituto Colombiano para la Evaluación de la Educación (Icfes) tiene dos frentes: (i) ofrecer servicios de evaluación de la educación en todos sus niveles, y (ii) fomentar la investigación sobre la calidad educativa para contribuir a su mejoramiento a partir de los datos de las pruebas aplicadas.

Para promover la investigación de la calidad educativa, el Icfes formuló el Programa de Investigación sobre Calidad de la Educación.

Una de las actividades centrales de este programa es el Seminario Internacional de Investigación sobre Calidad de la Educación, el cual fue concebido como un espacio académico anual para la presentación de estudios innovadores adelantados por investigadores nacionales e internacionales, además de los trabajos más destacados financiados por el Programa de Investigación sobre Calidad de la Educación en el marco de sus convocatorias anuales, dirigidas a estudiantes de posgrados y a grupos de investigación.

El seminario es un espacio académico anual para la presentación de estudios sobre la calidad educativa que busca ser un referente académico por presentar trabajos rigurosos, recientes e innovadores; un espacio de diálogo de investigadores de varias disciplinas, educadores de todos los niveles de educación, y diseñadores de política, interesados en el mejoramiento de la calidad educativa. Las presentaciones de las conferencias principales de los seminarios se encuentran disponibles en la página web del seminario <http://www.icfes.gov.co/index.php/component/k2/item/1751?Itemid=594>.

### *Objetivos*

El Seminario Anual de Investigación en Educación de llevó a cabo durante la semana del 3 al 6 de noviembre de 2015 en el hotel Marriott de Bogotá, sede calle 26. A diferencia de las ediciones anteriores, el evento fue temático, haciendo énfasis en los índices de calidad. Para 2016, el foco serán los factores asociados.

Sin embargo, se mantiene el propósito de dar a conocer ante la comunidad académica, docentes y hacedores de política algunas de las investigaciones más recientes que hacen uso de pruebas estandarizadas sobre logro académico; con la motivación adicional de exponer algunos de los proyectos de investigación adelantados por el Icfes.

### *Participantes*

Desde su primera versión, el seminario ha estado dirigido a investigadores, profesores y estudiantes de educación superior de diversas disciplinas, incluida educación, psicología, sociología, economía, estadística y otras ciencias sociales en las cuales exista interés en investigar sobre la calidad de la educación.

### *Programa académico, talleres y conferencias*

En el seminario de 2015, los talleres no se realizaron simultáneamente con las conferencias internacionales y nacionales, estos se desarrollaron los días 3 y 4 de noviembre, mientras que las conferencias y la exposición de afiches tuvieron lugar el 5 y 6 de noviembre.

Las conferencias internacionales se enfocaron en la presentación de estudios acerca de los índices de calidad, el valor agregado y la rendición de cuentas (*accountability*). Como en años anteriores, los criterios para la selección de estudios de las conferencias principales fueron los siguientes:

- Estudios de investigación aplicada que sugieran ideas para mejorar la calidad de la educación, bien sea en el aula o al nivel de políticas públicas.
- Estudios que empleen pruebas estandarizadas de logro académico, como las pruebas Saber.
- Estudios recientes que hagan énfasis y busquen responder preguntas actuales, enriqueciendo la discusión académica del país.
- Estudios multidisciplinares con diversidad en preguntas y métodos de investigación.
- Estudios con heterogeneidad de preguntas según el nivel educativo.
- Estudios que permitan conocer experiencias de sistemas educativos de diferentes países, con espacio permanente para países de América Latina.
- Las conferencias principales cuentan con la revisión previa de pares académicos.

Los talleres fueron de carácter introductorio, enfocados en cuatro temas: análisis estadístico de relaciones causales, modelos estructurales usando MPLUS, evaluación de impacto y uso de base de datos de factores asociados a las pruebas Saber. Como se mencionó anteriormente, estos se desarrollaron dos días antes del inicio de la conferencia con el objetivo de que los participantes logaran asistir a ambas actividades.

En la versión 2015 del seminario se mantuvo la figura del comentarista, introducido en el seminario de 2013, y cuyo fin es abordar los resultados de cada investigación en el contexto nacional. Actores nacionales de alto nivel tuvieron la oportunidad en cada una de las sesiones de presentar un comentario de 15 minutos que debía vincular los trabajos presentados en el módulo comentado con el contexto colombiano, haciendo énfasis en las implicaciones de las investigaciones más que en la metodología empleada. En las tablas 1 a 3 se muestra los temas tratados en el evento.

**Tabla 1.** Conferencias principales

Conferencista	Conferencia
Susanna Loeb, Center for Education Policy Analysis, Stanford University	Autonomy, decision making and quality in education
Morgan S. Polikoff, University of Southern California	Instructional alignment under no child left behind
Thomas Ahn, University of Kentucky	The impact of no child left behind's accountability sanctions on school performance. Regression discontinuity evidence from North Carolina
Jesse M. Cuhna, Naval Postgraduate School	Measuring value added in higher education
Peter Tymms, Durham University	The first seven years at school
Juan Esteban Carranza, Banco de la República, sede Cali	Nuevos programas y calidad de la educación superior en Colombia
Ernesto San Martín, Pontificia Universidad Católica de Chile	Valor agregado en la educación superior

**Tabla 2.** Conferencias de índices de calidad y *Call for Papers*

Conferencista	Conferencia
<b>Índices de calidad</b>	
Silvia Montoya (Unesco)	Monitoring Learning for All: Challenges and Opportunities of Composite Indexes to serve the SDGs
Alexandre André dos Santos (director de Evaluación y Educación Básica del INEP, Brasil)	El IDEB y educación nacional. Retos y perspectivas
Silvana Godoy y Jorge Duarte (Subdirección de Análisis y Divulgación del Icfes)	Índice Sintético de Calidad Educativa
Anderson García	La experiencia de Medellín en la medición de la calidad educativa. El índice multidimensional de calidad educativa como herramienta para la toma de decisiones
Luis Alfredo Muñoz	Índice de la calidad educativa de Bogotá
Carlos Henríquez Calderón (secretario ejecutivo de la Agencia de Calidad de la Educación de Chile)	Sistema de calidad e indicadores de seguimiento para Chile
Claudia C. Rave (Gobernación de Antioquia)	Sistema de calidad e indicadores de seguimiento para Medellín

Jorge A. Palacio (Subdirección Académica IDEP)	Las voces de los sujetos. Sistema de monitoreo al Plan Sectorial de Educación de Bogotá
<b>Call for Papers</b>	
Gloria Bernal Nisperuza y Luz Karime Abadía (Icfes, convocatorias)	Brechas de género en el rendimiento escolar a lo largo de la distribución de puntajes. Evidencia pruebas Saber 11
Silvia Gómez Soler y Paula Herrera (Icfes, convocatorias)	¿Vale la pena prepararse? Efectos de la participación en actividades preparatorias sobre el desempeño en el examen Saber 11
George Dueñas	Predicción automática de la dificultad de preguntas abiertas de respuesta corta con propósitos educativos
Ramiro Rodríguez Revilla y Diana Carolina López Cuevas	Valor agregado en lengua extranjera de la educación superior en Colombia
Guillermo Augusto García Méndez	Factores asociados al desempeño de los estudiantes de psicología en las pruebas Saber Pro
Fabio Sánchez Torres y Glenda Palacio (Icfes, convocatorias)	Etnoeducación y desempeño escolar en la región pacífica colombiana
Jaime Andrés Sarmiento Espinel y Adriana Carolina Silva Arias (Icfes, convocatorias)	Desigualdad de oportunidades en la educación media y superior en Colombia
Darío Maldonado	Efectos de largo plazo del Programa Mamá Canguro sobre resultados educativos y de fuerza laboral
Jhon James Mora Rodríguez	Educación media y mano de obra industrial en Colombia. Inequidad y pobreza
Ana María Zamora Moreno (Icfes, convocatorias)	¿Qué pasa en el salón de clases con el examen Saber 11? Una etnografía escolar sobre colegios exitosos
Andrea Cuenca (Icfes, convocatorias)	Desigualdad de oportunidades en Colombia. Impacto del origen social sobre el desempeño académico y los ingresos de graduados universitarios
Osmar Leandro Loaiza Quintero y Guillermo David Hincapié Vélez	Convergencia espacial del rendimiento educativo en Colombia (2000-2012)

Tabla 3. Talleres.

Conferencista	Taller
Andrés Gutiérrez (director del Departamento de Estadística y Evaluación del Icfes)	Introducción al análisis estadístico de relaciones causales
Lina Saldarriaga (miembro del Comité Asesor del Icfes)	Introducción al uso de modelos estructurales usando MPLUS
Luis Carlos Reyes (profesor asistente de la Universidad Javeriana)	Introducción a la evaluación de impacto
Carolina Lopera, Sandra Roperó y Lorena Wilches (Subdirección Estadística del Icfes)	Uso de bases de datos de factores asociados a las pruebas Saber

## II

## Comparación de contenidos de los seminarios 2010 a 2015

Las tabla 4 y 5 muestran la distribución de las conferencias magistrales y alternas de los seis seminarios anuales realizados hasta el momento, según líneas de investigación. También se muestra la heterogeneidad de disciplinas de los estudios presentados. En cuanto a líneas de investigación (filas de la tabla), estas se organizan de investigación básica (las primeras filas) a investigación aplicada (las últimas filas). Así, en la tabla 4 hay tres grupos de líneas de investigación: metodología de investigación, evaluación de la educación y sistemas de rendición de cuentas que emplean datos de logro académico. En la tabla 5 se mantienen las líneas con excepción de la metodología. En cuanto a heterogeneidad de disciplinas de los estudios seleccionados, para las conferencias magistrales de los seminarios, estas se agruparon en “educación” (color azul, con 24 conferencias principales y 15 alternas en el período), “economistas” (color rosado, con 17 conferencias principales y 27 alternas en el período), investigadores vinculados al Icfes (color verde, con 10 conferencias principales y 13 alternas), “ciencias sociales” (color amarillo, con 2 conferencias principales y 5 alternas en el período), provenientes de “convocatorias” (color morado, con 13 conferencias alternas en el período) y “otras disciplinas” como ciencias, geografía, etc. (color café, con 9 conferencias alternas).

**Tabla 4.** Conferencias principales de los seminarios Icfes (2010-2015) por línea de investigación y disciplina.

Ejes temáticos	2010	2011	2012	2013	2014	2015
<b>Metodología</b>						
Metodología	Shavelson					
<b>Evaluaciones estandarizadas de logro</b>						
Medición y psicometría		Cervantes				
		Cuevas				
Factores asociados	Porta				Lopera	
	Fernandes					
Rendición de cuentas y uso de resultados			Koretz		Koretz	Loeb
			Ponczek			Polikoff
Valor agregado						Ahn
			San Martin	Molina	Melguizo	Cuhna
			Domingue	Wiley	Cuellar	Tymms
				Bellegem		Carranza
				Domingue		San Martin
Índices de calidad						Montoya
						Dos Santos
						Godoy
						García
						Muñoz
						Henríquez
						Rave
					Palacio	
<b>Intervenciones y logro</b>						
Colegios	Bonilla	Bellei				
Ciudadanía y desarrollo humano		Chaux	Roco			
		Uzaheta				
Currículo	Schmidt					
Equidad	Willms		Reardon		Madeira	
Género		Cabezas				
Maestros		Lavy		Martínez	Nomi	
				Meckes	Kleickmann	
Tiempo de estudio		Lavy(b)				
		Menezes				
Educación superior	Saavedra			Briggs		
				Urzúa		
				Arum		
Intervenciones tempranas				Figlio	Rosales	

**Tabla 5.** Conferencias alternas de los seminarios Icfes (2010-2015) por línea de investigación y disciplina.

Ejes temáticos	2010	2011	2012	2013	2014	2015
<b>Evaluación</b>						
Evaluación		Cuchimaque		Ramírez	Garzón	Zamora
		Bohórquez		Sánchez	Márquez	Dueñas
Medición				Mariño		
				Gaertner		
Psicometría	Cervantes	Cuevas		Antonio	Herrera et al.	
	Cuevas	Herrera		Gamboa		
Factores asociados	Uzaheta ©	Silva	Montes	Cotto	López	García
	García			Olivera	Orobio	
	Arana			Cortés	Orejuela	
	Corsi				Córdoba	
	Fontecha				Equipo CESAC	
Equidad						Sarmiento
						Mora
						Cuenca
						Loaiza
Género					Bernal	
Pre-Icfes					Gómez	
Rendición de cuentas y uso de resultados				Zamora		
				Navarro		
Valor agregado		Zevallos		Alonso		Ramírez
		Bogoya				
		Alonso				
<b>Intervenciones y logro</b>						
Intervenciones y logro	Valderrama			Garcés		Maldonado
				Botello		
Colegios						
Ciudadanía y desarrollo humano	Uzaheta (b)	Chaux				
					Cristancho	
					Vargas	
Currículo		Aragón		Rojas		
				Orozco		
Equidad		Duarte		Montoya		
		Celis				
Género				Mendoza		
Maestros	Ome			Cardona		
Pares	Martínez			Breton	Quiñones	
Educación superior		Barón		Acosta	Maldonado	
		Gil		Sánchez	Melo	
		Gómez		Virgüez	Ospina	
		León		Cortés	Bayona	
		Solano				

Pre-Icfes				Niño		
Eficiencia escolar					Mendoza	
					Rodríguez	
					Godoy	
Efecto escuela					López	
					Alonso	
					Virgüez	

### III

## Las presentaciones académicas

En este capítulo se presentan las principales ideas de las conferencias. Sus presentaciones al público se encuentran en la página web del Icfes. Estos *resúmenes no pretenden reemplazar la lectura de libros y los estudios de los autores*, sino transmitir las ideas más relevantes planteadas en el evento.

### Conferencias sobre rendición de cuentas

#### *Autonomía, toma de decisiones y calidad de la educación*

Sussana Loeb, Center for Education Policy Analysis, Stanford University

El eje central de la conferencia es el uso de las pruebas estandarizadas a nivel de escuela en Estados Unidos, resaltando su utilidad para los hacedores de política, quienes deben alinear sus metas a los objetivos propuestos por el sistema educativo e identificar las necesidades que este posee para tomar decisiones de la forma más acertada.

En la década de los 80, en Estados Unidos, surge la preocupación acerca de la efectividad del sistema educativo en razón al bajo desempeño, evidenciando que regular las credenciales de los profesores no aseguraba la calidad de la enseñanza y sugiriendo que quizás más rendición de cuentas podría mejorar el desempeño de los colegios.

Loeb muestra cuatro formas para que las escuelas puedan mejorar: (i) regulación estatal, donde el gobierno central fija las reglas; (ii) proceso político, en el cual se deben producir personas con habilidades y alineación a las metas para la toma de decisiones; (iii) mecanismos de mercado, como la posibilidad de la elección de instituciones educativas; y (iv) una rendición de cuentas basada en resultados o procesos enfocados –a través de la implementación de exámenes a los estudiante– en analizar el desempeño del colegio. La ley “Que los niños no se queden atrás” de 2002 es el ejemplo más representativo de este tipo de sistemas.

En esta década también surge la necesidad de un sistema de rendición de cuentas a nivel de las escuelas para medir el desempeño del sector público y como respuesta a la baja efectividad del sistema educativo. En 1995 y bajo argumentos similares –bajo desempeño y lograr

mejoras en calidad educativa–, se desarrollaron las reformas con el objetivo de identificar un grupo de estándares de desempeño claro, medible y ambicioso para los estudiantes en un número de materias y para alinear el currículo a estos estándares.

En este contexto, la rendición de cuentas a nivel de colegios provee un mecanismo poderoso para superar, a través de incentivos, el problema de información. En el lenguaje de los economistas, principal-agente, en el cual los hacedores de política, padres y locales se enfrentan con la dificultad de monitorear las actividades de las escuelas, razón por la que podría existir un comportamiento opuesto por parte de los educadores al esperado por los agentes interesados en el buen desarrollo del sistema educativo. La medición y reporte del desempeño de las instituciones educativas acompañado de consecuencias positivas o negativas al lograr o no el nivel de desempeño propuesto motiva a los educadores a concentrarse en las áreas y materiales que están siendo medidos. Esto, a su vez, permite a los hacedores de política evaluar que tan exitoso ha sido la escuela en obtener las metas propuestas por el Estado.

El diseño de un sistema de rendición de cuentas puede tener un impacto significativo en la naturaleza y la fortaleza de los incentivos que enfrentan las escuelas para incrementar el logro de los estudiantes en las asignaturas evaluadas. Los sistemas de rendición de cuentas pueden tomar dos aproximaciones cuando se consideran las consecuencias de su implementación: (i) la inclusión de sanciones o recompensas explícitas para las escuelas que fallan o cumplen las expectativas, y (ii) la inclusión frecuente de sanciones de las instituciones educativas que no logran el objetivo de desempeño.

El alcance y dominio de los indicadores de rendición de cuentas difieren tanto en número como en tipo de pruebas, por lo que el gobierno central enfrenta disyuntivas importantes al determinar qué tan amplia debe ser la base del sistema implementado; en particular, establecer un sistema que alinee la rendición de cuentas a un pequeño grupo de resultados tiende a limitar el alcance de la educación provista a estudiantes. De otro lado, un amplio grupo hace más difícil medir de manera confiable y puede confundir el enfoque de las escuelas y el personal del distrito. Un factor limitante para el alcance y el dominio de la rendición de cuentas es la tecnología disponible para la evaluación del progreso de los estudiantes, además de estar atenuado por los retos que impone cada materia para su evaluación.

Loeb hace énfasis en dos aproximaciones de la medición del desempeño a nivel de colegio con base en los puntajes de evaluación: las medidas de “estado”, que juzgan el logro de los estudiantes dado un nivel de desempeño; y las medidas de “crecimiento”, a menudo llamadas medidas de “valor agregado”, que evalúan las escuelas sobre el grado en cual los estudiantes mejoran en el desempeño de su evaluación de un año a otro.

Las medidas de estado y de valor agregado miden resultados diferentes: las primeras se enfocan en el porcentaje de estudiantes que obtienen el nivel de desempeño que busca

incentivar a la escuela a incrementar los niveles de rendimiento, además es atractivo para los hacedores de política porque establece el mismo objetivo para todo los grupos de estudiantes y hace que las instituciones educativas presten atención a los estudiantes de bajo rendimiento, quienes en el pasado pueden no haber recibido la atención necesaria. En contraste, los modelos de “crecimiento” incentivan a los colegios a mejorar el desempeño de los estudiantes independientemente del nivel absoluto que se logró. Bajo cualquier aproximación, las medidas de valor agregado o de estado, los errores en la medición del desempeño del estudiante pueden generar rankings inconsistentes en el tiempo, lo que implica que estas medidas no son perfectas.

Otro aspecto relevante para quienes diseñan los sistemas de rendición de cuentas es determinar los estudiantes a tener en cuenta cuando se evalúa su aprendizaje, pues hay diversos incentivos para excluir estudiantes de la evaluación, como los niños que no tienen un desempeño que genere una buena valoración de la institución educativa.

Uno de los propósitos de los sistemas de rendición de cuentas podría ser prestar atención a los grupos de estudiantes que tradicionalmente presentan bajo desempeño, en particular, y establecer ciertos niveles de desempeño para estos subgrupos de estudiantes dentro de una población escolar. La política “Que los niños no se queden atrás” en Estados Unidos hace hincapié no solo en el rendimiento académico del cuerpo estudiantil, sino en subgrupos de estudiantes definidos por su raza, ingreso económico y discapacidad.

El periodo de tiempo en el cual se debe rendir cuentas es también un tema relevante en el diseño del sistema. Incrementar el periodo del tiempo en el cual las escuelas son evaluadas reduce los errores de medición y su incorrecta, aunque esto requiere más tiempo para la recolección de datos y la obtención de índices de mejora o empeoramiento.

Claro está que los sistemas de rendición de cuentas podrían no generar un mayor logro en el desempeño de los estudiantes. Uno de los argumentos para justificar esta afirmación es el cambio de los temas que se enseñan por parte de los docentes, al ajustar el contenido expuesto en clase a los temas del examen, lo cual podría no ser una preocupación si las evaluaciones cubren un amplio rango del material importante para la sociedad. En contraste a las evaluaciones con altas implicaciones para los estudiantes, las pruebas con bajas implicaciones tienen dos desventajas: la primera hace referencia a que los estudiantes podían tomar en serio este tipo de pruebas, implicando un incremento de los errores de medida, y la segunda es el no alineamiento a los estándares valorados por los hacedores de política.

Un factor adicional que podría hacer que un sistema de rendición de cuentas no produzca los efectos esperados en el nivel de desempeño de los estudiantes es el *comportamiento estratégico*; es decir, las instituciones educativas pueden adoptar estrategias que artificialmente mejoran los puntajes del grupo de estudiantes en la asignatura evaluada. Por ejemplo, la asignación selectiva de estudiantes a programas de educación especial.

Finalmente Loeb resume las ventajas de los sistemas de rendición de cuentas, entre las que se destacan los debates públicos con metas claras, el establecimiento de medidas consistentes que permita comparar y trabajar con profesores y colegios, y el establecer un sistema más amplio que reduzca la necesidad de regulaciones que son a menudo poco efectivas. Igualmente señala algunas desventajas, como el enfoque excesivamente ajustado, los errores de medición que pueden crear incentivos erróneos y la desviación de las metas locales.

### *La alineación de la enseñanza bajo la política "Qué ningún niño se quede atrás"*

Morgan Polikoff, University of Southern California

La conferencia busca explicar la importancia que tiene estudiar la enseñanza, la cual se entiende como las alineaciones de las instrucciones entre las instituciones y su importancia en la valoración de los resultados de los estudiantes. La enseñanza tiene tres dimensiones claves: (i) contenido, ¿qué se enseña? (ii) forma y calidad, ¿cómo se enseña?, y (iii) el tiempo de enseñanza, cuya medición y conceptualización es una meta relevante para el desarrollo de investigaciones y el diseño de políticas públicas.

Polikoff hace énfasis en la primera dimensión a través del concepto de alineación, señalado como un elemento fundamental de un sistema de rendición de cuentas, porque se desea que las medidas se alineen con las metas y porque permite observar el desempeño de los docentes por medio de un currículo contra el cual se mide el desempeño de los estudiantes. La alineación es el grado en el que el contenido de las asignaturas concuerda con los estándares y las evaluaciones propuestas, igualmente el conferencista propone una definición de *contenido* como la intersección de temas y niveles de demanda cognitiva.

Al establecer estándares y evaluaciones, la alineación surge como una variable que modera el efecto entre el aprendizaje y los esquemas establecidos, con el objetivo de motivar a docentes a enseñar el contenido especificado en los documentos normativos (estándares, pruebas, materiales curriculares). La medición de la alineación de la instrucción puede ser difícil, en particular si queremos medir la alineación del rendimiento de un año ajustado a un grupo de estándares; es decir, cuánto se ajustó lo que enseñaron los docentes a un plan de estudios o a una evaluación.

La observación puede no ser factible al no poder monitorear lo que ocurre en el aula para contrastar si se está alineado o no al contenido propuesto en el currículo. Adicionalmente, el desarrollo de las encuestas también puede representar una dificultad para la medición de la alineación en razón a la mala formulación de las preguntas, como aquellas encuestas donde el docente es cuestionado directamente acerca de su conocimiento sobre el concepto de alineación, donde el margen de error está representado por la poca claridad que pueda tener el profesor sobre la definición de alineación y como se realiza en su clase.

Polikoff presentó los resultados de tres estudios, en los cuales realiza un análisis del contenido por medio de las encuestas de currículos representativos, una herramienta para medir el contenido de lo que enseñan los docentes, así como el contenido producido por investigadores de la Universidad de Michigan. Los datos para enseñanza y contenido están agrupados en matrices de temas por niveles de demanda cognitiva. Para cada asignatura hay una lista amplia de tópicos analizados. Por ejemplo, en matemáticas, una temática son las fracciones y se pregunta sobre si el estudiante memoriza o aprende de una forma conceptual (tabla 6).

El lenguaje del contenido de la encuesta ha sido desarrollado y revisado por expertos para estructurar un lenguaje neutro que permita compararlo con la instrucción que los docentes reportan estar enseñando y haciendo uso. Es decir, se ha puesto este documento en términos de lo que se espera del desempeño de los estudiantes y se les pide a los profesores un reporte de lo que están enseñando para poder hacer un comparativo.

La alineación puede ser medida a través de un índice que hace uso de los datos de la encuesta de currículo representativo, donde  $x_i$  es la proporción de contenido en la celda  $i$  de la matriz  $x$  (la matriz que representa la enseñanza del profesor) y  $y_i$  es la proporción de contenido en la celda  $i$  de la matriz  $y$  (matriz que representa el contenido de los estándares establecidos o una evaluación):

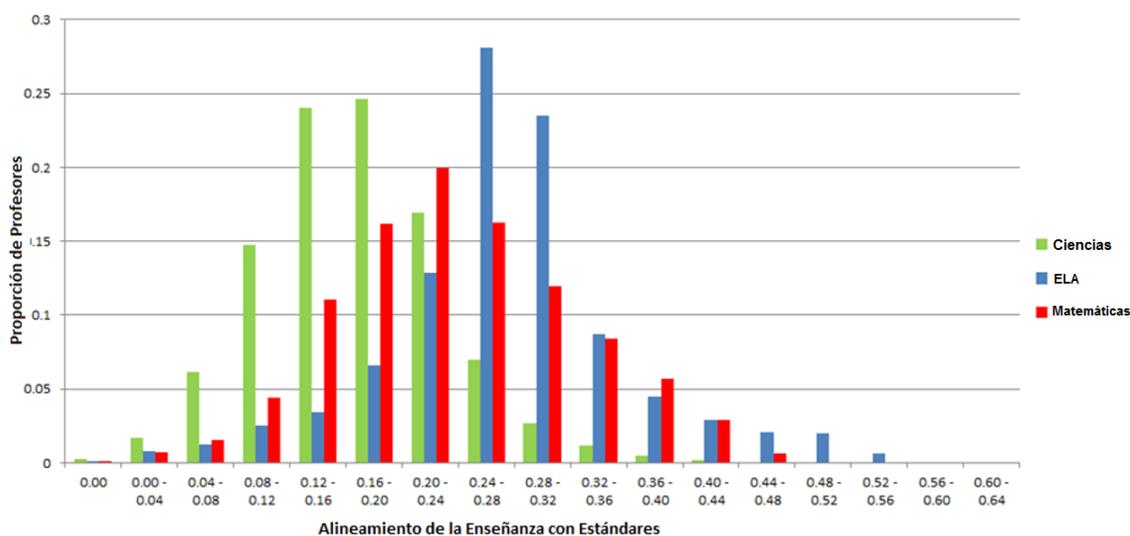
$$\text{Alineación} = 1 - \frac{\sum_i |x_i - y_i|}{2}$$

La variable de alineación está entre 0-1, siendo 0 perfecto desalineamiento de la instrucción y 1 alineación perfecta de la enseñanza, siendo los valores entre 0 y 1 los que representan la proporción de contenido común. Polikoff presenta tres estudios cuya variable de resultado es el índice expuesto anteriormente; el primero de ellos hace referencia al grado de alineación en las escuelas de Estados Unidos y si está ha incrementado. En este artículo se hace uso de datos provenientes de una encuesta de aproximadamente 3.000 profesores, distribuidos en estados, grados y asignaturas diferentes, recolectados entre 2002-2011, y trata de responder cómo la instrucción se alinea a estándares y evaluaciones. También muestra el cambio de la alineación de la enseñanza en los años que la ley federal "Que los niños no se queden atrás" estuvo en vigencia.

**Tabla 6.** Medición del contenido mediante encuestas de currículos representativos.

Tiempo en el Tema		Grados K-12 Temáticas de Matemáticas			Expectativas para los Estudiantes en Matemáticas			
<Ninguno>	*	Sentido Numérico/Propiedades/Relaciones	Memorizar Hechos/Definiciones/Fórmulas	Rendimiento de Procedimientos	Demuestra Comprensión de Ideas matemáticas	Conjetura/Generaliza/Prueba	Soluciona problemas no rutinario/Realiza Conexiones	
⊙ ① ② ③	101	Notación Posicional	⊙ ① ② ③	⊙ ① ② ③	⊙ ① ② ③	⊙ ① ② ③	⊙ ① ② ③	
⊙ ① ② ③	102	Números Enteros	⊙ ① ② ③	⊙ ① ② ③	⊙ ① ② ③	⊙ ① ② ③	⊙ ① ② ③	
⊙ ① ② ③	103	Operaciones	⊙ ① ② ③	⊙ ① ② ③	⊙ ① ② ③	⊙ ① ② ③	⊙ ① ② ③	
⊙ ① ② ③	104	Fracciones	⊙ ① ② ③	⊙ ① ② ③	⊙ ① ② ③	⊙ ① ② ③	⊙ ① ② ③	

**Figura 1.** ¿Qué tan alineada esta la enseñanza?



En la figura 1 se observa que las cifras en promedio son más bajas de lo que se esperaría, cada proporción de profesores tiene un grado de alineación inferior al 50%; pero en contraste se puede evidenciar una buena distribución, lo que permite explorar factores asociados a este tipo de alineación. Lo observado permite concluir que la alineación es más baja de lo que se percibía cuando se le pregunta a los docentes como se estaban alineando.

En relación a la alineación de la enseñanza en los años que la ley “Que los niños no se queden atrás” estuvo en vigencia, el índice de alineación es regresado sobre el año, controlando por la composición de la muestra con efectos fijos de estado y grado para evaluar si el alineamiento de la instrucción ha ido incrementándose año tras año. El coeficiente de la variable año representa el cambio promedio en un año en la alineación dentro de estados y grados, promediado entre estados y grados:

$$\text{Alineación}_{ijk} = \beta_0 + \beta_1 \text{Año} + u_j + f_k + \varepsilon_{ijk}$$

La variable *Año* es el año de realización de la encuesta de currículo representativo,  $u_j$  es un grupo de efectos fijos de estado y  $f_k$  es un grupo de efectos fijos por grado. Este análisis se hace por separado para cada asignatura, señalando como principal resultado el mayor grado de alineación en matemáticas respecto lenguaje y ciencias, al igual que un incremento de la alineación en cuanto a los estándares de contenido.

Haciendo uso de datos similares al estudio anterior, Polikoff estudia la puesta en práctica de la alineación de la enseñanza, resaltando su importancia en razón de establecer cuál debe ser la población objetivo para el desarrollo docente, como profesores nuevos que no se han alineado a los estándares o docentes que no tienen el grado de contenido necesario para dictar la asignatura. El análisis se realiza a nivel de escuela, utilizando las variables de experiencia y educación del profesorado como predictores focales (años de experiencia de enseñanza, título de maestría, contenido del área de enseñanza), adicionalmente se controla por características de los estudiantes.

Entre los principales hallazgos de este estudio se encuentran que los docentes que poseen entre 1 y 11 años de experiencia están más alineados respecto a quienes no tienen experiencia, un curso con mayor contenido está asociado con una mayor alineación de la enseñanza, así como docentes que realizaban más cursos, especialmente en matemáticas, tienden a estar más alineados a los estándares respecto a quienes no realizan este tipo de actividades.

Finalmente, el tercer estudio expuesto hace uso del trabajo de la Fundación Bill & Melinda Gates sobre medidas de enseñanza efectiva. Se trata de responder en qué grado la alineación de la enseñanza con el contenido de los estándares y las evaluaciones reportado por los profesores está asociado con el valor agregado al desempeño del estudiante, y hasta qué punto las medidas de calidad pedagógica moderan la relación de alineación instruccional con el valor agregado al logro estudiantil. La muestra está compuesta por 327 profesores entre 4° y 8° grado de matemáticas e inglés pertenecientes a seis distritos escolares urbanos en seis estados de Estados Unidos. Los docentes completaron la encuesta de currículo representativo al final del año escolar y se tenían los contenidos de los estándares estatales y las evaluaciones para observar cómo estaban alineados, además de las medidas de valor agregado y de observación de docente calculadas para el estudio realizado a mayor escala.

Polikoff señala el hallazgo de correlaciones pequeñas (entre 0.1 y 0.2) entre alineación y valor agregado significativas, las cuales eran tan grandes como las medidas de observación de la calidad pedagógica con valor agregado. En razón a que el estudio se llevó a cabo en seis estados, fue posible evidenciar que la relación entre el valor agregado y las medidas de instrucción era bastante débil en algunos estados respecto a otros, lo que el conferencista denominó la sensibilidad de las pruebas frente a la instrucción.

*El impacto de las sanciones de la política: “Qué los niños no se queden atrás” en el desempeño de la escuela. Evidencia de una regresión de discontinuidad para Carolina del Norte*

Thomas Ahn, University of Kentucky

La conferencia explica la política “Que ningún niño se quede atrás” y expone el estudio formulado por Ahn y Vigdor que busca evaluar el impacto de las sanciones impuestas por esta ley a las escuelas a través de la evaluación del desempeño de los estudiantes.

La ley se implementó con un alto apoyo bipartidista en 2002, con la meta de tener un conocimiento perfecto de matemáticas y lectura para 2014, requiriendo el uso de evaluaciones estandarizadas para el desempeño de los estudiantes para acceder a los fondos federales. Los legisladores observaron la imposibilidad de lograr el 100% de desempeño en un año, lo que condujo a los estados a establecer un nivel mínimo de desempeño para las escuelas.

La política “Que ningún niño se quede atrás” estuvo limitada por ser un programa federal. No fue un modelo de crecimiento ya que todos los estados tenían que tener la misma disponibilidad de datos y algunos no estaban siguiendo el desempeño de los estudiantes año a año, razón por la que fue un modelo novedoso. Además el programa fue impulsado por castigos ad hoc que no se sabía si tendrían un claro efecto en el desempeño de las instituciones (tabla 7).

En la actualidad, la política ha llegado a ser impopular y genera desdén en los sindicatos de profesores, pues se ha percibido que daña a las escuelas que necesitan más ayuda y a algunos padres no les gusta su régimen de evaluación. Ahn señala algunos hallazgos previos del estudio como la efectividad de algunas sanciones y la ineffectividad de otras, y las grandes mejoras producto del cambio de liderazgo, el cual está asociado con más comunicación entre la facultad y la administración, un ambiente de respeto y una evaluación más justa de los profesores.

**Tabla 7.** Sistema de sanciones de la política “Que ningún niño se quede atrás”.

Régimen de sanción de la política: “NCLB”	
Años de Pérdidas en el PAA	Sanción
1	Ninguna, ubicación en lista de vigilancia, desarrollo de un plan de mejoramiento
2	Transferencia con servicio de transporte a escuelas de más alto rendimiento.
3	Servicios de educación suplementario. (Sólo estudiantes FRL)
4	“Sanciones Correctivas”
5	Formulación de un plan de reestructuración
6	Implementación del programa de reestructuración

En justificación a que las escuelas sancionadas y las no sancionadas son bastante diferentes, no es posible establecer una comparación entre estas. La política “Que ningún niño se quede atrás” fue una ley nacional aplicada una vez, por lo que no es un experimento natural. Así, la solución metodológica propuesta (técnica econométrica) que se puede emplear es la “regresión de discontinuidad” debido a que el régimen de sanciones impuesto trata a las escuelas de forma muy diferente si el desempeño de un subgrupo de estudiantes cae en el lado cualquiera de un umbral.

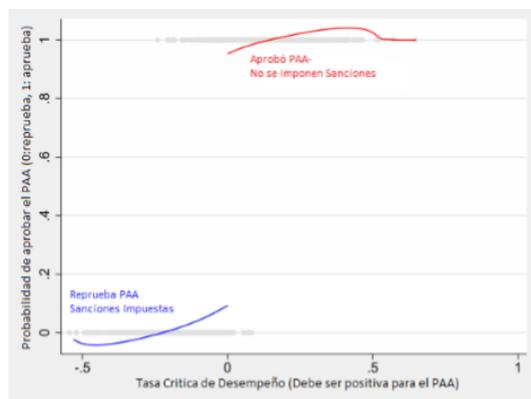
La tasa de desempeño del colegio determina si la institución va a ser castigada o no, con base en los estudiantes que hayan alcanzado el puntaje del progreso adecuado anual (AYP, por sus siglas en inglés). En la figura 2 se puede observar que la tasa de desempeño crítico es 0 y que quienes se encuentran por debajo, representados por la línea azul, no logran el nivel de desempeño esperado, lo que conlleva a las instituciones educativas a estar expuestas a las sanciones de la política, en contraste a las escuelas cuyo nivel de desempeño está en un nivel superior al punto de corte y no se ven expuestas a castigos.

La regresión de discontinuidad permite estimar el impacto de las sanciones de la política. El grupo de comparación son las instituciones educativas que están a la izquierda cerca del punto de corte, puesto que son instituciones educativas idénticas a aquellos colegios justo a la derecha del punto cero, pero que por alguna razón como explica Ahn no pudieron alcanzar el nivel mínimo de desempeño.

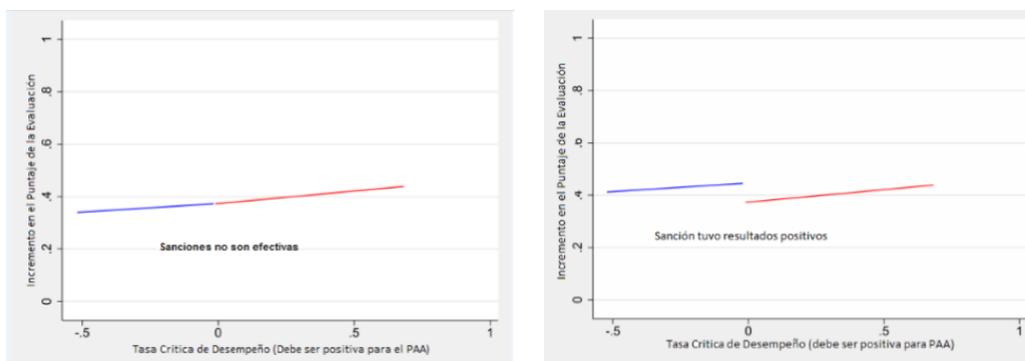
Recordando que la línea roja representa las escuelas que no están expuestas a las sanciones de la política y la azul a aquellas que están sancionadas, se puede observar que las sanciones no son efectivas si no existe un cambio para las instituciones educativas expuestas al régimen de castigos en contraste a las que no fueron reprendidas. En cambio, si los colegios sancionados (línea azul) cambian respecto a los colegios que no estuvieron expuestos a sanciones, se puede afirmar que las sanciones produjeron resultados positivos. Los resultados que se presentan es una estimación de la brecha (figura 3) entre la línea roja y la línea azul.

Ahn señala que los resultados estadísticamente significativos se reflejan para matemáticas al evaluar el efecto de no lograr el puntaje del *progreso adecuado anual*: los colegios sancionados mejoraron su puntaje en aproximadamente en 0.2 desviaciones estándar, lo cual es un resultado modesto y está acorde con la literatura previa al estudio (tabla 8 y figura 4).

**Figura 2.** Regresión de discontinuidad.

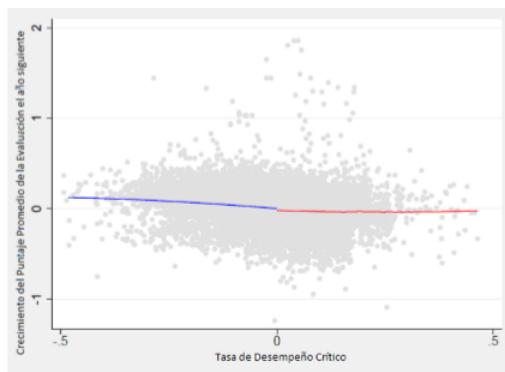


**Figura 3.** Identificación efectos RD.



**Tabla 8.** Efecto de fallar en el PAA.

	Lectura (n=8264)	Matemáticas (n=8266)
En el ancho de banda ópt.	-0.00002	0.0180*
En la mitad del ancho de banda ópt.	0.0061	-0.0117
Doble ancho de banda ópt.	-0.0025	0.0282***
En el ancho de banda óptimo, incluyendo transferencias	-0.0004	0.0220***

**Figura 4.** Estimación de fallar en el PAA.

Teniendo en cuenta el historial de rendición de cuentas del colegio, se divide la muestra y se evalúa inicialmente la exposición de la institución educativa a una advertencia a ser sancionado por primera vez, es decir, colegios no sancionados previamente. Se encontró un efecto positivo: en promedio, las escuelas aumentaban su desempeño en aproximadamente 4 o 5 desviaciones estándar en el año siguiente (tabla 9 y figura 5).

En caso de que los colegios no logren el PAA de manera reiterativa, serán expuestos a las sanciones subsecuentes, como el desarrollo de un plan de mejoramiento, la transferencia de estudiantes a mejores colegios o servicios suplementarios de educación (tutorías) sin tener un efecto relevante en el desempeño de las escuelas. Y al evaluar el impacto de sanciones más fuertes el estudio señala que la amenaza de un plan de reestructuración en las escuelas tiene un impacto positivo y significativo (tabla 10).

Una de las preocupaciones del establecimiento del sistema de incentivos por la ley era crear la motivación para enfocarse en cierto subgrupo de estudiantes y que esto derivara en afectaciones en el desempeño de otros. Para tratar este tema, los autores dividen a los estudiantes en cuartiles de desempeño, excluyendo a quienes presentan buenos resultados, esto en justificación de que la política establece que el 70% de los estudiantes deberían superar el mínimo nivel de desempeño establecido y estos no requerirían un nivel de esfuerzo muy alto para superar el punto de corte. Al observar los cuartiles más bajos y los más altos, ninguno de estos grupos fueron afectados por las sanciones impuestas. Ahn cree que una posible explicación a este hecho es la utilización de dos sistemas de rendición de cuentas, a nivel federal y estatal, los cuales actúan de forma complementaria para que ningún niño se quede atrás.

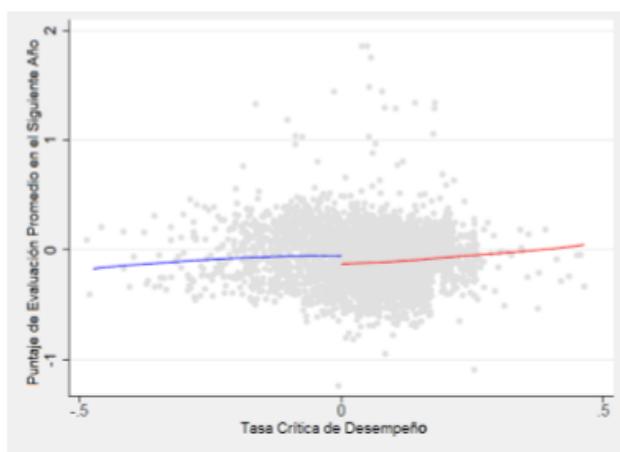
En concordancia con el análisis de los castigos, como por ejemplo la reestructuración y el cambio en la dirección de los colegios, se concluyó que traen efectos positivos sobre el desempeño de los colegios pero no se logró identificar a través de qué mecanismo.

En función de descubrir lo que ocurre cuando hay cambios en el liderazgo, se usa la encuesta de condiciones laborales con el objetivo de conectar los resultados de los estudios de la rendición de cuentas con la percepción de los profesores sobre el liderazgo, una vez las instituciones han sido sometidas a sanciones.

**Tabla 9.** Efecto de estar expuesto a la primera sanción.

	Lectura (n=3241)	Matemáticas (n=3241)
En el ancho de banda óptimo	0.0203	0.0473**
En la mitad del ancho de banda óptimo	0.0086	-0.0197
2 veces el ancho de banda óptimo	0.0200*	0.0.0681***
En el ancho de banda óptimo, (incluyendo transferencias)	-0.0004	0.0759***

**Figura 5.** Efecto de fallar en el PAA para escuelas no sancionadas previamente.



**Tabla 10.** RD estima el impacto de sanciones mayores.

Acción Correctiva, amenaza de plan de reestructuración	Lectura (n=549)	Matemáticas (n=549)
En el ancho de banda óptimo	0.035	0.0231
Plan de Reestructuración, amenaza de reestructuración	Lectura (n=442)	Matemáticas (n=442)
En el ancho de banda óptimo	-0.0558**	0.0278
En medio ancho de banda	-0.052	0.0486
En el doble del ancho de banda	-0.0937*	0.0197
Reestructuración	Lectura (n=367)	Matemáticas (n=367)
En el ancho de banda óptimo	0.0264*	0.0547**
En medio ancho de banda	0.0562***	0.0664*
En el doble del ancho de banda	0.0205	0.0609**

**Tabla 11.** El liderazgo realmente importa.

Muestra	Liderazgo Efectivo	Mutuamente Respetuoso	Liderazgo Apoyo a Profesores
Muestra Completa	0.0122	0.0132	-0.0021
Amenaza solamente	-0.0332	-0.0194	<b>-0.510*</b>
Sanción Baja	0.0174	0.0109	0.0005
Sanción Alta	<b>0.0905**</b>	<b>0.1290**</b>	<b>0.1096**</b>

**Tabla 12.** Los profesores se sienten evaluados más justamente.

Muestra	Evaluación apropiada	Evaluación Consistente	Profesores Reciben Retroalimentación
Muestra Completa	0.0062	0.0003	0.0068
Amenaza solamente	-0.0086	-0.0399	-0.0375
Sanción Baja	-0.0276	-0.0363	0.0005
Sanción Alta	<b>0.1061**</b>	<b>0.0984**</b>	<b>0.0990**</b>

En la tabla 11 se puede ver la proporción de profesores que están de acuerdo que los directores tienen un liderazgo efectivo. Hay un mutuo acuerdo entre la facultad y la administración sobre un liderazgo que apoya a los profesores si se considera la muestra completa de colegios: aquellos que inicialmente han sido advertidos, quienes presentan una sanción baja o aquellos que se enfrentan a una sanción más estricta. Los resultados estadísticamente significativos se encuentran para las escuelas que enfrentan una sanción alta. Estos efectos se pueden interpretar como la proporción de profesores en una institución cuya sanción fue alta (cambio de líderes) que considera que el liderazgo efectivo ha aumentado en un 10%, el respeto mutuo entre administración y facultad ha aumentado en un 12.9% y el liderazgo que apoya a los docentes ha aumentado en un 12%.

Igualmente, la tabla 12 se puede entender como la proporción de profesores que están de acuerdo con que haya una evaluación apropiada y consistente, y que los docentes reciben retroalimentación considerando la muestra completa de colegios. Luego, se puede interpretar que la proporción de profesores en una institución cuya sanción fue alta (cambio de líderes) considera que hay un aumento de 10% en promedio en la consistencia de las evaluaciones, retroalimentación docente y evaluación apropiada.

Se puede observar también que cuando hay una alta rotación de líderes, no se está implementando un cambio estructural o hay una gran asignación de recursos, lo que está ocurriendo es un cambio en la relación entre los funcionarios administrativos y los docentes que conlleva a un mejor desempeño en las pruebas. Ahn concluye afirmando que contrario a la opinión popular, se ha podido evidenciar que las sanciones impuestas por la política pueden funcionar, generando una llamada de alerta a las escuelas que no tienen el rendimiento

esperado y mejorando el desempeño de las instituciones educativas, en algunos casos a través de cambios drásticos en el liderazgo o a través de factores como el incremento de la comunicación, un ambiente de respeto y una evaluación más justa de los docentes.

## Conferencias sobre valor agregado

### *Midiendo el valor agregado en educación superior*

Jesse M. Cuhna, Naval Postgraduate School

La conferencia habla sobre las mediciones de valor agrado en educación superior, su uso y las limitaciones a las que se enfrentan para su construcción. La relevancia de este tema se sustenta en los incentivos generados a los profesores a producir mejores resultados, donde el expositor contextualiza a la audiencia en Estados Unidos, país en el cual los estados están interesados en las medidas de crecimiento con el objetivo de desarrollar una mejor rendición de cuentas. Igualmente, Cuhna señala que la investigación previa se ha enfocado en educación primaria y secundaria, dejando una presión para la rendición de cuentas a nivel de educación superior, contexto en el que los investigadores encuentran el reto de desarrollar indicadores cuantificables de desempeño.

Investigaciones previas sugieren que las medidas de valor agregado pueden capturar el desempeño diferencial de instituciones y profesores, pero evidencian grandes diferencias al extrapolar el modelo empleado en educación primaria y secundaria a educación superior: (i) las evaluaciones estandarizadas año a año no son generalmente empleadas en educación superior, situación que conlleva a plantear como variables de resultado las tasas de persistencia, las tasas de graduación y los salarios después de haber culminado la universidad; (ii) no se puede aislar la influencia de factores específicos en el proceso educativo de la habilidad preexistente del estudiante con base en resultados observados una sola vez; (iii) existe un problema de selección en razón a que los estudiantes escogen deliberada y sistemáticamente su accesos a la universidad, lo cual dificulta identificar el efecto de asistir a la institución de educación superior separadamente del efecto de características preexistentes; y (iv), finalmente, los estudiantes universitarios especializan su instrucción, lo cual no puede ser medido de forma óptima con una evaluación estandarizada.

Cuhna señala la importancia de entender la educación superior como un proceso multidimensional derivado de objetivos como productividad económica, consumo de conocimiento e incluso, en ciertos contextos, puede considerarse objeto social, cuya cuantificación no puede ser fácil. Por ejemplo, si se piensa en el nivel de ingreso como un resultado de la productividad económica de ir a una universidad, entonces es posible afirmar que se genera más valor agregado si el asistir a una institución de educación superior le permite al individuo

obtener un ingreso más alto comparado a si hubiese asistido a otra universidad. Esta medida del proceso educativo puede presentar problemas en razón de factores no observables que hacen que el salario no sea una buena aproximación de una medida de crecimiento, reconociendo en las evaluaciones estandarizadas una herramienta para observar las habilidades y el conocimiento de una persona después de realizar sus estudios universitarios, como el conferencista resalta se ha hecho Colombia con Saber Pro, no obstante remarcando que Estados Unidos estas evaluaciones no se realizan.

Antes de explicar las metodologías empleadas en su estudio, Cuhna hace algunas acotaciones sobre la regresión de discontinuidad, haciendo énfasis que es una técnica econométrica que permite evaluar el efecto de asistir a una universidad específica; es decir, permite analizar el efecto para aquellos individuos que pudieron acceder a una institución de educación superior respecto a quienes no lo lograron, pero no obstante es una metodología que presenta ciertas limitaciones en cuanto a que compara solo los individuos cercanos al punto de corte. En el contexto colombiano, se hace uso de una regresión de discontinuidad basado en evaluaciones estandarizadas para medir el efecto de asistir a una mejor universidad sobre los puntajes, el empleo y los salarios.

La metodología propuesta por Cuhna y Miller (2014) busca estimar el valor agregado relativo de instituciones individuales por medio de una regresión a nivel de estudiante que explique el resultado de interés a través de: (i) las diferencias observables en las características de pre-matrícula del estudiante; (ii) las diferencias no observadas en las preferencias de los estudiantes capturadas por el perfil de aplicación y aceptación del estudiante; y (iii) los efectos fijos por universidad a la cual el estudiante se matriculó. En consecuencia, si se considera una función de producción educativa, podría estimarse las diferencias en los resultados de los estudiantes entre escuelas cuando el resultado de interés es observado. En Colombia, Cuhna señala que el lado izquierdo de la ecuación puede caracterizarse como Saber Pro y el lado derecho de la ecuación como Saber 11:

$$Y_{is} = \delta Y_{i,PRE} + \Phi X_{i,PRE} + \sum_{s=1}^n \beta_s E_s + \varepsilon_{is}$$

$E_s$  es un subgrupo de indicadores para los matriculados en varias universidades  $s$ ,  $Y_{is}$  es el nivel de ingreso para el estudiante  $i$ , quien asistió a la escuela  $s$ ,  $Y_{i,PRE}$  es el nivel de ingreso antes de matricularse y  $X_{i,PRE}$  es un vector de características observables del estudiante. Aunque lo más conveniente sería interpretar  $\beta_s$  como un promedio a nivel de escuela, se puede dar una confusión en la forma de analizar este coeficiente como una relación causal a partir de varios argumentos: (i) los estudiantes pueden escoger las universidades de manera que son inobservables y están correlacionados con el resultado de interés, y (ii) la medida de resultado, en este

caso el nivel de ingreso, antes de matricularse puede ser no observado, lo que conlleva a estimar de la siguiente forma:

$$Y_{is} = \Phi X_{i,PRE} + \sum_{s=1}^n \beta_s E_s + \varepsilon_{is}$$

$X_{i,PRE}$  capturaría los efectos de selección no observables que pueden estar correlacionados con los resultados y la elección de la escuela. Además, la falta de información antes de que el estudiante se matricule implica que  $\beta_s$  refleje las diferencias relativas en los resultados entre escuelas. En relación con lo anterior, Cuhna expone la necesidad del uso de datos longitudinales que reflejen el proceso educativo antes, durante y después de la vida universitaria. Así, se hace uso de los datos disponibles a nivel individual provenientes de universidades (créditos tomados, graduación, aplicaciones, admisiones, aceptaciones, etc.), las escuelas secundarias (graduación, cursos tomados, puntajes de las evaluaciones, programas, etc.), las comisiones del seguro de desempleo (ganancias trimestrales) y la evaluación de los servicios del sector privado (puntajes SAT y encuestas de respuestas).

El ejercicio empírico mostrado en la conferencia emplea datos de Texas. La muestra está comprendida por los graduados de las escuelas públicas en cinco cohortes (1998-2002), se incluyen a quienes se matricularon en una universidad pública de Texas hasta 2010 y se excluyen los estudiantes que decidieron no ir a la universidad o aquellos provenientes de otros estados que estudian en instituciones de educación superior de Texas. A través de estos datos se definen varias medidas de resultados: (i) los salarios, la suma de las cuatro ganancias trimestrales en el octavo año calendario después de haber terminado la escuela secundaria; (ii) la persistencia del segundo año de universidad, es decir, lograr acreditar 30 horas en una universidad pública o privada en un programa de cuatro años en el final del sexto año académico después de graduarse de la secundaria; y (iii) la graduación, recibir un diploma de una universidad pública o privada en Texas al final del sexto año académico después de graduarse del colegio.

Examinando cómo cambian los salarios controlando sucesivamente por grupos de covarianzas antes de matricularse a la universidad, puede mostrarse en la tabla 13 estimaciones de  $\beta_s$ , donde la variable de resultado es el logaritmo de las ganancias ocho años después de haberse graduado del colegio. La columna 1 incluye efectos fijos por año de graduación, los cuales absorben algunas diferencias agregadas entre años, como efectos de inflación o choques macroeconómicos.

Cuhna hace énfasis en alguno de los resultados eliminando algunas de las filas de la tabla 13, explicando que los coeficientes se pueden interpretar como las diferencias porcentuales promedio en los salarios de los matriculados en varias universidades de Texas, relativo a los matriculados en Texas A&M. Por ejemplo, los matriculados en UT Dallas ganan 12% menos que los matriculados en Texas A&M. En las columnas progresivas se incluyen más controles,

disminuyendo el problema de selección ya nombrado y agregando efectos fijos de género y raza, con lo que se puede observar que no hay cambios considerables. Luego se adicionó información sobre la escuela secundaria a la que asistió el estudiante, en la columna subsiguiente se agregó la prueba SAT (prueba parecida a Saber 11) y por último se controló por efectos fijos de grupo de aplicación. Lo anterior permite ver una reducción de los coeficientes; es decir, hay menos diferencias entre las universidades comparadas respecto a Texas A&M.

Los administradores o los padres de familia desean saber si las diferencias anteriormente mencionadas existen, si son estadísticamente significativas, lo cual se realiza a través de una comparación entre las diferentes universidades por medio del test de Pairwise, como muestra la tabla 14, resaltando la limitación de comparar entre grandes grupos de colegios.

Enfocándonos en otros resultados, como la tasas de persistencia y de graduación e incluyendo controles de forma subsecuente, se puede observar que las diferencias desaparecen, además de destacar la relevancia de otros temas como la omisión de las ganancias de quienes no están trabajando en el sector formal. La figura 6 muestra las relaciones entre las medidas de valor agregado y se observa la correlación entre el crecimiento generado por (i) persistencia y ganancias, (ii) la tasa de graduación y las ganancias, y (iii) la persistencia y la graduación. Para concluir, se señala que las regresiones *kitchen-sink* son probablemente la mejor forma de medir valor agregado entre universidades con los datos disponibles. Igualmente, Cunha remarca la importancia acerca de que las medidas de valor agregado no generen incentivos perversos a los educadores y den información útil para todos los actores del sistema.

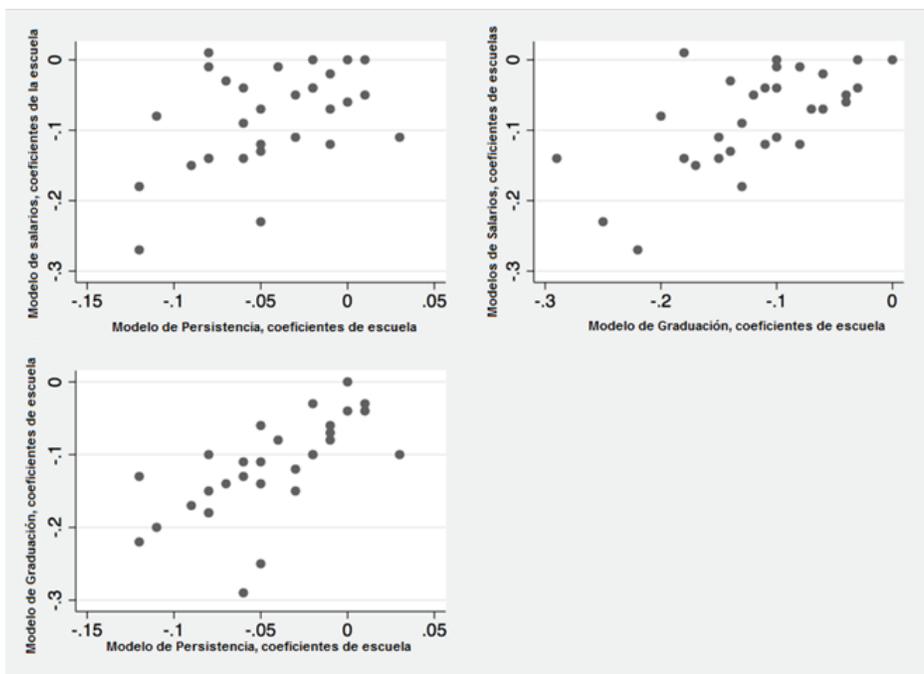
**Tabla 13.** Regresiones de salarios sobre indicadores de matrícula.

Outcome =	ln(earnings) (1)	ln(earnings) (2)	ln(earnings) (3)	ln(earnings) (4)	ln(earnings) (5)
UT Tyler	-0.24** (0.04)	-0.23** (0.04)	-0.15** (0.04)	-0.13** (0.04)	0.01 (0.08)
Texas Women's U.	-0.30** (0.03)	-0.21** (0.03)	-0.12** (0.03)	-0.08** (0.03)	0.00 (0.05)
Texas A&M U.	-	-	-	-	-
Texas Tech U.	-0.18** (0.01)	-0.18** (0.01)	-0.07** (0.01)	-0.05** (0.01)	0.00 (0.03)
S. F. Austin State U.	-0.28** (0.01)	-0.25** (0.01)	-0.16** (0.01)	-0.12** (0.01)	-0.01 (0.03)
Tarleton State U.	-0.28** (0.02)	-0.29** (0.02)	-0.14** (0.02)	-0.10** (0.02)	-0.01 (0.05)
UT Dallas	-0.12** (0.02)	-0.12** (0.02)	-0.14** (0.02)	-0.11** (0.02)	-0.02 (0.05)
West Texas A&M U.	-0.33** (0.02)	-0.32** (0.02)	-0.13** (0.03)	-0.09** (0.03)	-0.03 (0.06)
U. of North Texas	-0.31** (0.01)	-0.28** (0.01)	-0.20** (0.01)	-0.16** (0.01)	-0.04 (0.03)
UT Pan American	-0.52** (0.01)	-0.47** (0.02)	-0.19** (0.03)	-0.14** (0.03)	-0.04 (0.03)
TAMU Galveston	-0.21** (0.04)	-0.21** (0.04)	-0.17** (0.04)	-0.13** (0.04)	-0.04 (0.07)
TAMU Commerce	-0.36** (0.02)	-0.33** (0.02)	-0.25** (0.03)	-0.20** (0.03)	-0.05 (0.05)
Texas State U.	-0.24** (0.01)	-0.23** (0.01)	-0.13** (0.01)	-0.10** (0.01)	-0.05 (0.03)
UT Austin	-0.13** (0.01)	-0.12** (0.01)	-0.12** (0.01)	-0.11** (0.01)	-0.06** (0.02)
TAMU Corpus Christi	-0.31** (0.02)	-0.29** (0.02)	-0.17** (0.02)	-0.13** (0.02)	-0.07 (0.04)
S. Houston State U.	-0.29** (0.02)	-0.25** (0.02)	-0.18** (0.02)	-0.14** (0.02)	-0.07 (0.03)
Lamar U.	-0.35** (0.02)	-0.32** (0.02)	-0.21** (0.03)	-0.16** (0.03)	-0.08 (0.04)
UT Arlington	-0.25** (0.02)	-0.22** (0.02)	-0.17** (0.02)	-0.14** (0.02)	-0.09* (0.04)
UT Permian Basin	-0.30** (0.04)	-0.28** (0.04)	-0.13** (0.05)	-0.08 (0.05)	-0.11 (0.12)
U. of Houston	-0.26** (0.01)	-0.21** (0.01)	-0.20** (0.01)	-0.17** (0.01)	-0.11** (0.03)
Prairie View A&M U.	-0.59** (0.02)	-0.41** (0.02)	-0.28** (0.02)	-0.23** (0.02)	-0.12** (0.04)
TAMU Kingsville	-0.50** (0.02)	-0.46** (0.02)	-0.23** (0.03)	-0.18** (0.03)	-0.12** (0.04)
Midwestern State U.	-0.42** (0.02)	-0.40** (0.02)	-0.23** (0.03)	-0.18** (0.03)	-0.13** (0.05)
Angelo State U.	-0.43** (0.02)	-0.41** (0.02)	-0.20** (0.02)	-0.16** (0.02)	-0.14** (0.04)
UT San Antonio	-0.40** (0.01)	-0.37** (0.01)	-0.24** (0.02)	-0.19** (0.02)	-0.14** (0.03)
U. of Houston Dwtm.	-0.42** (0.02)	-0.33** (0.02)	-0.22** (0.02)	-0.17** (0.02)	-0.14** (0.04)
Sul Ross State U.	-0.48** (0.04)	-0.45** (0.04)	-0.19** (0.04)	-0.14** (0.04)	-0.15 (0.08)
TAMU International	-0.41** (0.03)	-0.36** (0.03)	-0.26** (0.05)	-0.20** (0.05)	-0.18** (0.06)
Texas Southern U.	-0.78** (0.02)	-0.60** (0.02)	-0.43** (0.03)	-0.37** (0.03)	-0.23** (0.05)
UT El Paso	-0.63** (0.02)	-0.58** (0.02)	-0.40** (0.03)	-0.35** (0.03)	-0.27** (0.05)
Year FE	Yes	Yes	Yes	Yes	Yes
Race FE, gender		Yes	Yes	Yes	Yes
H.S. FE, H.S. demographics			Yes	Yes	Yes
SAT score, SAT demographics				Yes	Yes
Application group FE					Yes
Observations	169,239	169,239	169,239	169,239	169,239
Adjusted R <sup>2</sup>	0.03	0.03	0.05	0.06	0.06

**Tabla 14.** Prueba de Pairwise p-valores. Resultado = ln (salarios).

	UT Tyler	Texas Women's U.	Texas A&M U.	Texas Tech U.	S. F. Austin State U.	Tarleton State U.	UT Dallas	West Texas A&M U.	U. of North Texas	UT Pan American	TAMU Galveston	TAMU Commerce	Texas State U.	UT Austin	TAMU Corpus Christi	S. Houston State U.	Lamar U.	UT Arlington	UT Permian Basin	U. of Houston	Prairie View A&M U.	TAMU Kingsville	Midwestern State U.	Angelo State U.	UT San Antonio	U. of Houston Dwrtn.	Sul Ross State U.	TAMU International	Texas Southern U.	UT El Paso		
UT Tyler	.99																															
Texas Women's U.		.99																														
Texas A&M U.			.99																													
Texas Tech U.				.99																												
S. F. Austin State U.					.99																											
Tarleton State U.						.99																										
UT Dallas							.99																									
West Texas A&M U.								.99																								
U. of North Texas									.99																							
UT Pan American										.99																						
TAMU Galveston											.99																					
TAMU Commerce												.99																				
Texas State U.													.99																			
UT Austin														.99																		
TAMU Corpus Christi															.99																	
S. Houston State U.																.99																
Lamar U.																	.99															
UT Arlington																		.99														
UT Permian Basin																			.99													
U. of Houston																				.99												
Prairie View A&M U.																					.99											
TAMU Kingsville																						.99										
Midwestern State U.																							.99									
Angelo State U.																								.99								
UT San Antonio																									.99							
U. of Houston Dwrtn.																										.99						
Sul Ross State U.																											.99					
TAMU International																												.99				
Texas Southern U.																													.99			
UT El Paso																														.99		

**Figura 6.** Relaciones entre medidas de valor agregado.



### *Los primeros siete años en la escuela*

Peter Tymms, Durham University

El estudio expuesto aborda el impacto a largo plazo del acceso de los niños a educación de alta calidad en las escuelas durante varios años sucesivos. La motivación de este estudio se deriva de los hallazgos de la investigación de Sanders y Rivers: “un niño expuesto a tres años de enseñanza de un gran profesor o un mal profesor genera un efecto importante en un plazo de tres años, y los efectos sobre el logro del estudiante son aditivos y acumulativos con poca evidencia de efectos compensatorios”. En el contexto inglés, los niños comienzan por ley el colegio a la edad de 5 años, aunque los padres los envían a una clase de preescolar un año antes. La investigación se basa en los datos recolectados de 73.347 alumnos, quienes comenzaron la escuela en Inglaterra en 1999 y se les hizo seguimiento hasta 2005, siendo evaluados en al menos dos ocasiones. Es importante señalar que las instituciones educativas no tienen la obligación de usar las evaluaciones provistas por el instituto no gubernamental (Centro para la Evaluación y el Monitoreo, CEM), del cual Tymms hace parte, sin embargo pagan por ser monitoreadas.

Antes de entrar en el detalle de su estudio, Tymms expone la forma estatal u oficial de llevar a cabo el monitoreo sobre las escuelas, donde la información de las pruebas es publicada en periódicos, como es común en Inglaterra, ejerciendo presión sobre las instituciones educativas. Luego, si un colegio no tiene un buen resultado en el ranking publicado, es posible que la institución trate de ocultar el problema, pues se envía un inspector a que observe el desempeño de la institución y genere un informe de sus deficiencias.

Lo anterior ha tenido consecuencias desastrosas, como suicidios en el profesorado, razón por la que el conferencista sugiere que el monitoreo debería suministrar información para los docentes y no para agentes ajenos al sistema. Así los docentes tendrían información sobre los niños, permitiéndoles encontrar los problemas (niños con problemas de lectura y/o rezagados en matemáticas) y buscar soluciones.

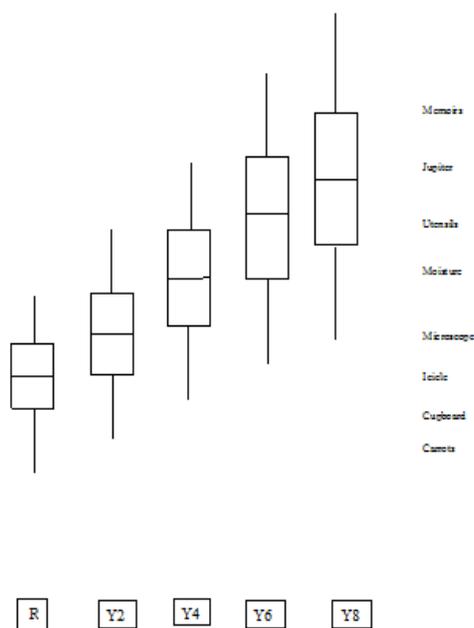
El diseño de evaluación es un reto, particularmente para los niños de cuatro años a los que no se les pueden aplicar evaluaciones en papel ya que no saben leer, por lo que se les hace una secuencia de preguntas durante 20 minutos para atenuar la dificultad de los lapsos cortos de concentración y dar confiabilidad a la prueba. En estas se hace hincapié en tres áreas del conocimiento: vocabulario, lectura y matemáticas. Tymms da ejemplos del tipo de preguntas al que podría enfrentarse un infante, como las asociaciones de imágenes con palabras para el vocabulario, el número de palabra leídas e identificación de números para las matemáticas.

Con respecto a vocabulario, hay datos disponibles de 1.140 escuelas y se emplea modelos multinivel para su análisis, en los cuales se estima las correlaciones entre los puntajes de los alumnos y el puntaje promedio de las escuelas año a año, además de usar las medidas obtenidas en años previos para “explicar” las mediciones en años subsecuentes.

Observando la figura 7, en el eje horizontal tenemos la temporalidad escolar, con niños que empiezan desde la etapa más temprana (preescolar) y se observan cada dos años, mientras que en el eje vertical se encuentra el vocabulario. Es importante señalar que los puntajes crecen con el tiempo, pero si examinamos el puntaje más alto pareciese que crece de manera estable, en contraste a los puntajes bajos, donde esto no aplica, pues crece pero la curva se está aplanando. Esta característica es muy importante y se relaciona con el nivel de lectura y matemáticas. Adicionalmente se puede ver que los niños de cuatro años más capacitados tienen un vocabulario más desarrollado respecto a los menos capacitados y que están terminando la primaria, implicando un gran reto para el docente en este año.

Tymms analiza los coeficientes de correlación para vocabulario como ejemplo y señala que si observamos los más débiles, estos se encuentran aproximadamente en 0.3 al analizar el bagaje en lectura, pero al realizar una prueba de matemáticas y de lectura se encuentran coeficientes de 0.7. El investigador señala que las correlaciones más altas que se han podido evidenciar últimamente se observan en los países de la OECD, en donde Colombia no tiene el mejor rol hasta el momento. Entre los estudiantes y los colegios observados en relación a los puntajes de vocabulario al inicio, tenemos que la correlación en promedio es 0.52 y que se esperaría que desapareciera con el tiempo, pero es una cifra bastante constante (tabla 15).

**Figura 7.** Crecimiento en el área de vocabulario.



Si miramos el progreso entre un colegio y otro (tabla 16), va a existir un puntaje dentro del primer año correspondiente al 21% de variación, lo cual es bastante alto como medida de avance de una institución educativa. También es factible observar que al incrementarse los años, el valor agregado va disminuyendo. Tymms señala que estudios previos muestran que las estimaciones deben estar entre 5% y 15%, y resalta que no es posible evidenciar una relación entre el progreso que se alcanza en un año, tanto a nivel de infantes como de escuelas, con el progreso alcanzado al año siguiente.

En la figura 8 se analizan los efectos de los impulsos repetidos en varios años en vocabulario, lectura y matemáticas. Si en vocabulario se experimenta un impulso en el año 1 seguido de un impulso en el año 2, el estudiante tendría un 0.2 de desviación estándar. Los impulsos pueden ser aditivos año a año, y un impulso en el año inicial seguido por impulsos en los tres años seguidos resulta en una ventaja educativa de cerca de un tercio de una desviación estándar en los años siguientes.

Observando qué ocurre en la asignatura de matemáticas, Tymms señala la falta de datos en justificación a la realización de una prueba nacional de lectura y matemáticas. Como en vocabulario, se usaron modelos para cuantificar qué ocurriría a los estudiantes que habían experimentado varios niveles de valor agregado en diferentes años. Al igual que el análisis previo con vocabulario, el impulso importante está en el primer año, por lo que la ventaja del estudiante promedio está cerca de un quinto de desviación estándar en su educación primaria. Respecto a lectura hay efectos similares, pero hay pequeñas diferencias, como que las correlaciones entre estudiantes año a año fueron más altas como lo las correlaciones a nivel de escuela.

Para resumir, los mayores estímulos se dieron en los primeros años y se evidenciaron ventajas de su efecto a largo plazo. Estas ventajas son acumulativas, pues generan avances año tras año, y aditivas, lo que significa que intervenciones repetidas pueden incrementar el desempeño de los estudiantes año tras año. Finalmente, Tymms concluye destacando algunas observaciones relevantes de política: (i) el progreso año tras año es vital, (ii) un impulso tardío podría no ser eficiente, (iii) los niños en un contexto de pobreza requieren un esfuerzo continuo (impulsos repetitivos), y (iv) los profesores son claves en la obtención de un proceso educativo.

**Tabla 15.** Promedio entre estudiantes.

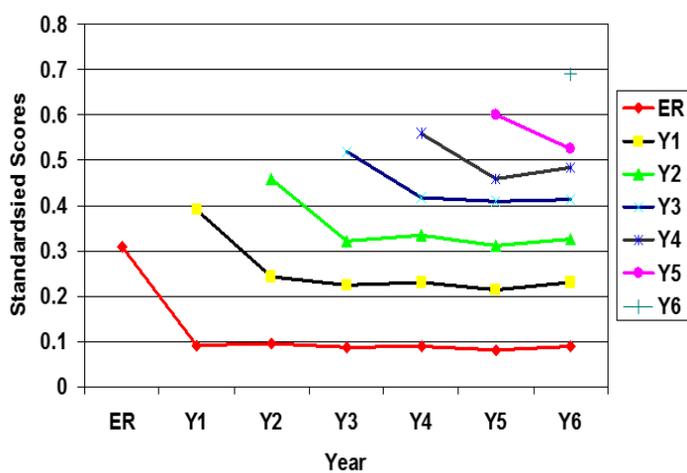
Año 1	Año2	Año 3	Año 4	Año 5	Año 6
0.52	0.54	0.53	0.52	0.50	0.50

**Tabla 16.** Promedio entre colegios.

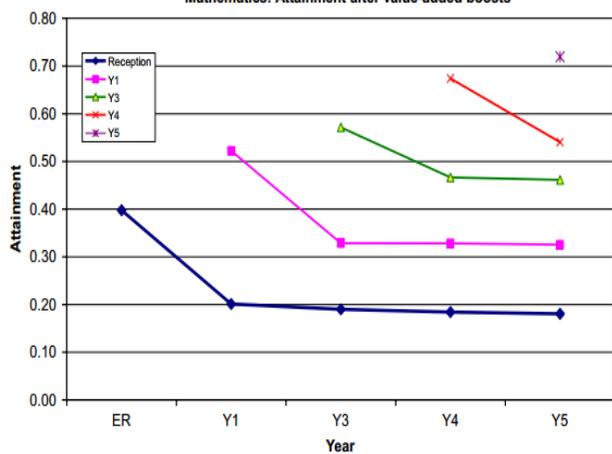
Año 1	Año2	Año 3	Año 4	Año 5	Año 6
0.60	0.67	0.67	0.66	0.66	0.62

**Figura 8.** Análisis de efectos de los impulsos en vocabulario, lectura y matemáticas.

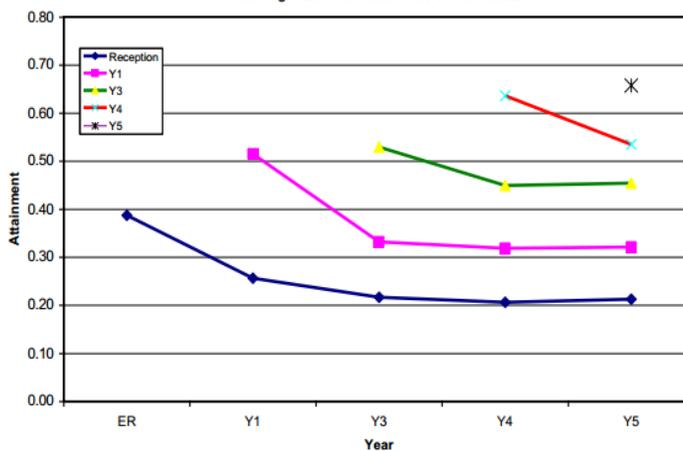
**Repeated Boosts: Vocabulary**



**Mathematics: Attainment after value-added boosts**



**Reading: Attainment after value-added boosts**



## *Nuevos programas y calidad de la educación superior en Colombia*

Juan Esteban Carranza, Banco de la República, sede Cali\*

El estudio trata acerca del crecimiento de la cobertura de educación superior en Colombia, el cual ha sido espectacular en los últimos quince años. En gran medida la calidad de las instituciones de educación superior depende de la calidad de estudiantes que salen de los colegios; es decir, el grueso del valor agregado de la educación superior está explicado por las características con las que llegan los estudiantes a la universidad. El objetivo de la investigación no es abordar solamente los números en el sentido de la cantidad, sino analizar el tipo de estudiantes que se están incorporando al sistema.

Entre los años 2000 y 2013 las matrículas del sistema de educación superior casi se duplicaron, especialmente en las instituciones con programas de duración de más de cuatro años. Este crecimiento se dio de dos formas: (i) al crecer los programas existentes y (ii) al crear nuevos programas. En este contexto, la calidad de la educación superior dependió del tipo de estudiante que atrajeron los programas universitarios, haciendo hincapié en dos características específicas: el ingreso de los hogares de los que provienen los estudiantes y la habilidad medida a través de las pruebas Saber 11.

La figura 9 muestra el crecimiento de la matrícula en el sector público y privado entre los años 2000 y 2013, evidenciado que en el sector no oficial pasó de 500.000 estudiantes a más de 900.000 y la matrícula del sector oficial creció aproximadamente de 300.000 a alrededor de 660.000. Igualmente, si se observa el tipo de programa, es posible ver en la figura 10 el incremento de los programas universitarios de cuatro años relativo a los de dos años de duración en el periodo anteriormente mencionado.

En la figura 11 se muestra el crecimiento de cuatro tipos de programa: (i) el área azul es la matrícula a los programas privados existentes, (ii) el área roja es la matrícula a los programas de universidades públicas existentes, (iii) el área verde es la matrícula a los programas privados nuevos y (iv) el área morada es la matrícula a los programas públicos nuevos. El grueso del crecimiento está explicado por el crecimiento del área verde y morada, que como se explicará adelante, atraen a gente distinta a los programas antiguos.

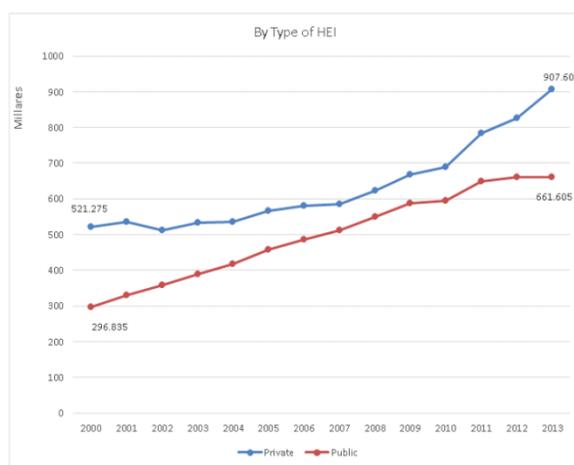
Pero se puede ser más específico, pues los programas nuevos pueden ser parecidos a los existentes. Carranza explica que en términos de la taxonomía del Icfes hay unos núcleos dentro de los cuales existen unos programas específicos; un ejemplo de ello puede ser administración de empresas, cuyo programa puede tener unas credenciales en una institución universitaria

---

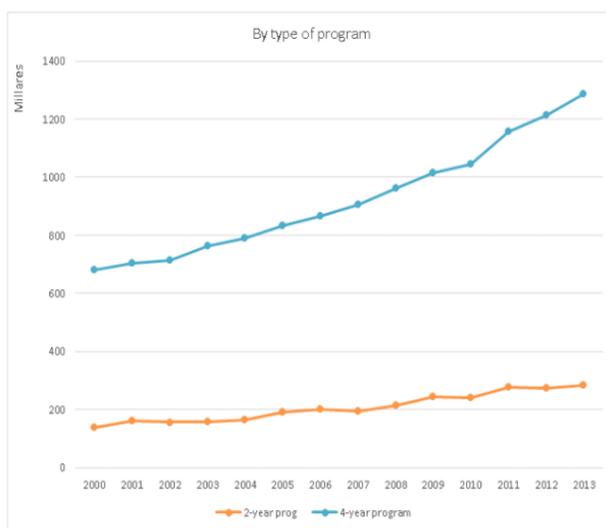
\* El estudio expuesto no representa una opinión de la Junta del Banco de la República o de su gerente general. Es un trabajo netamente académico de investigación.

pero puede ofrecer un nuevo programa dentro del campo de la administración. En la figura 12 se puede evidenciar que el área verde fue el de mayor crecimiento, es decir, programas privados nuevos en núcleos viejos.

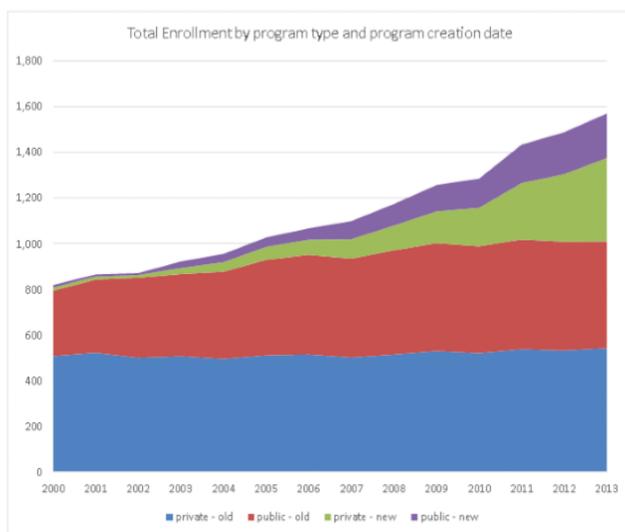
**Figura 9.** Crecimiento de la matrícula en el sector público y privado entre el año 2000 y 2013.



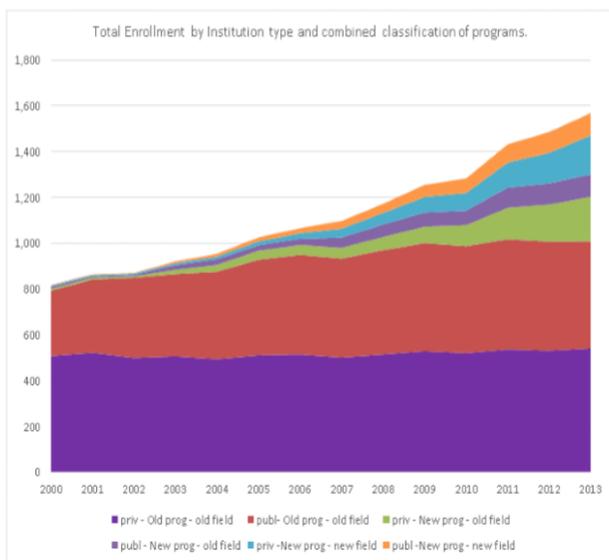
**Figura 10.** Crecimiento de la matrícula por duración del programa entre el año 2000 y 2013.



**Figura 11.** Crecimiento de los programas existentes y nuevos en universidades públicas y privadas.



**Figura 12.** Crecimiento de los programas de áreas tradicionales y nuevas en universidades públicas y privadas.



Para observar cómo ha variado la composición de los estudiantes que han entrado a la universidad, se toma primero el número de estudiantes de la cohorte que presentó las pruebas Saber 11 en 2000 y fueron a alguna institución de educación superior entre el 2000-2004, y luego el número de estudiantes de la cohorte que presentó las pruebas Saber 11 en 2009 y fueron a alguna institución de educación superior entre el 2010-2015. Lo anterior se muestra en las tablas 17 y 18, que se pueden entender de la siguiente forma: hacia la derecha el estudiante tiene un mayor nivel de habilidad, mientras las filas más baja representa el de menos nivel de ingreso posible. Al examinar esta tablas es posible evidenciar que ha existido un aumento de la matrícula de los estudiantes con un nivel alto de habilidad y con un nivel de ingreso más bajo.

Al realizar una comparación de las características de los estudiantes que entraron al sistema de educación superior en 2000 versus a los que ingresaron en 2009, se podrá examinar los que ingresaron a programas existentes como los que accedieron a programas nuevos. Carranza realiza un análisis para los programas de cuatros, años como se muestra en la tabla 19. La primera fila son los promedios en las pruebas Saber 11 y cada columna representa un tipo de programa. En 2000 solo tenemos programas viejos por definición, en contraste a 2009 donde se encuentran núcleos con nuevos campos o nuevos programas. Se puede evidenciar que el promedio en Saber 11 es similar para quienes accedieron a los programas ya existentes en 2000 como en 2009; pero para quienes ingresaron a nuevos programas, el promedio en la prueba estandarizada es menor, lo cual refleja un cambio en el tipo de estudiante que está atrayendo los programas nuevos respecto a los viejos.

**Tabla 17.** Distribución de la población que presentó Saber 11 en 2000 y va a una IES en 2001-2005, por ingresos y habilidad.

		Ability Quintile					Total
		1	2	3	4	5	
5+ SMLV	Abs.	916	1,419	2,364	4,125	11,146	19,970
	Rel (%)	0.77	1.19	1.98	3.46	9.35	16.76
3-5 SMLV	Abs.	1,394	2,255	3,289	5,015	8,895	20,848
	Rel (%)	1.17	1.89	2.76	4.21	7.46	17.49
2-3 SMLV	Abs.	2,521	3,755	5,049	7,019	9,722	28,066
	Rel (%)	2.12	3.15	4.24	5.89	8.16	23.55
1-2 SMLV	Abs.	4,257	5,925	7,563	9,638	11,382	38,765
	Rel (%)	3.57	4.97	6.35	8.09	9.55	32.53
<1 SMLV	Abs.	1,751	2,066	2,331	2,722	2,658	11,528
	Rel (%)	1.47	1.73	1.96	2.28	2.23	9.67
Total	Abs.	10,839	15,420	20,596	28,519	43,803	119,177
	Rel (%)	9	12.94	17.28	23.93	36.75	100

Source: Saber 11 and SPADIES.

**Tabla 18.** Distribución de la población que presentó Saber 11 en 2009 y va a una IES en 2010-2014, por ingresos y habilidad.

		Ability Quintile					Total
		1	2	3	4	5	
5+ SMLV	Abs.	701	1,288	2,176	4,149	12,912	21,226
	Rel (%)	0.32	0.58	0.98	1.88	5.84	9.6
3-5 SMLV	Abs.	1,624	2,770	4,277	6,903	12,630	28,204
	Rel (%)	0.73	1.25	1.94	3.12	5.71	12.76
2-3 SMLV	Abs.	4,017	6,169	8,554	12,334	16,364	47,438
	Rel (%)	1.82	2.79	3.87	5.58	7.4	21.46
1-2 SMLV	Abs.	10,061	14,329	18,118	22,427	23,556	88,491
	Rel (%)	4.55	6.48	8.2	10.15	10.66	40.04
<1 SMLV	Abs.	5,227	6,752	7,830	8,567	7,288	35,664
	Rel (%)	2.36	3.05	3.54	3.88	3.3	16.14
Total	Abs.	21,630	31,308	40,955	54,380	72,750	221,023
	Rel (%)	10	14.17	18.53	24.6	32.92	100

Source: Saber 11 and SPADIES.

Otro hecho visible recae en que los programas nuevos están atrayendo a los estudiantes más pobres en promedio, a gente un poco mayor respecto a los programas viejos, a más mujeres y a personas cuyas madres tienen un nivel menor de educación.

A través del uso de un modelo Logit, Carranza desea evaluar cómo han cambiado las características de los estudiantes entre 2000 y 2009 manteniendo todo lo demás constante. De forma intuitiva, las estrellas quieren decir que hay una correlación estadísticamente significativa; es decir, hay un resultado robusto y se tienen distintos regresores para controlar por diferentes variables. Las tablas 20 y 21 muestran una regresión logística en la que se observa cómo son las personas que eligen ir a la universidad respecto a las que deciden no ir y muestra algunos resultados aparentemente obvios como que la correlación entre presentar la prueba Saber 11 y la probabilidad de ir a la universidad es positiva. La segunda fila examina si dicha probabilidad se incrementó después de 2009 versus 2000, que da como resultado una correlación positiva.

Pero lo interesante es observar cómo cambiaron estas correlaciones entre los años 2000 y 2009, para lo cual se incluyó un término de interacción, es decir, se midió qué tanto se movió esa interacción en el periodo analizado. En 2009 se puede observar que la prueba Saber 11 facilita más el acceso a la universidad respecto a 2002, además se visualiza que las cifras de ingreso no muestran un patrón claro, lo que significa que este perdió importancia en justificación de como se explicó anteriormente: más personas de escasos recursos están accediendo a la universidad. La educación materna se vuelve más relevante en estos diez años, de igual forma las mujeres asisten más a la universidad pero se hace menos relevante para ingresar a la universidad. Esto puede verse en las tablas 22 y 23.

**Tabla 19.** Programas de cuatro años.

Tipo de programa: programa de 4 años

	Cohort: 2000		Cohort: 2009			T tests	
	Old prog	Old prog	New prog	New prog old field	New prog new field	New vs Old 2009	Before vs after - Old
	(1) mean	(2) mean	(3) mean	(4) mean	(5) mean	(2)-(3) Diff	(1)-(2) Diff
Average Sb 11 score std	0.595	0.596	0.249	0.146	0.302	0.347***	-0.001
Fam income =5+ SMLV	0.203	0.134	0.067	0.065	0.068	0.067***	0.069***
Fam income =3-5 SMLV	0.189	0.15	0.118	0.108	0.123	0.032***	0.039***
Fam income =2-3 SMLV	0.23	0.22	0.213	0.192	0.224	0.006***	0.010***
Fam income =1-2 SMLV	0.292	0.361	0.421	0.428	0.418	-0.061***	-0.068***
Fam income =<1 SMLV	0.086	0.136	0.18	0.206	0.167	-0.044***	-0.050***
Age	16.714	16.977	17.085	17.194	17.028	-0.107***	-0.263***
Is male	0.474	0.459	0.437	0.385	0.464	0.022***	0.014***
Mother ed: Primary	0.258	0.166	0.226	0.266	0.205	-0.059***	0.092***
Mother ed: secondary	0.318	0.427	0.468	0.462	0.471	-0.040***	-0.109***
Mother ed: T&T	0.201	0.146	0.132	0.121	0.138	0.014***	0.055***
Mother ed: University+ Works (=1)	0.223	0.261	0.175	0.151	0.187	0.086***	-0.037***
	0.041	0.047	0.057	0.063	0.054	-0.010***	-0.006***
Observations	86,081	126,158	36,074	12,325	23,749	162,232	212,239

Source: Saber 11 and SPADES. \*\*\*, \*\*, \* corresponds to 1%, 5% and 10% of significance, respectively.

**Tabla 20.** Coeficientes estimados de la regresión Logit. Variable dependiente: "¿Está en una IES?".

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Own sb11 score (std)	0.569*** (0.004)	0.657*** (0.004)	0.449*** (0.005)	0.475*** (0.005)	0.477*** (0.005)	0.472*** (0.005)	0.470*** (0.005)
After (=1 cohort 2009)	0.849*** (0.012)	0.557*** (0.011)	0.884*** (0.012)	2.721*** (0.068)	2.653*** (0.069)	2.391*** (0.070)	2.654*** (0.077)
Fam income: 1-2 mw	0.674*** (0.012)		0.481*** (0.012)	0.525*** (0.012)	0.543*** (0.012)	0.542*** (0.012)	0.533*** (0.012)
Fam income: 2-3 mw	1.194*** (0.013)		0.839*** (0.013)	0.910*** (0.013)	0.928*** (0.014)	0.939*** (0.014)	0.928*** (0.014)
Fam income: 3-5 mw	1.616*** (0.014)		1.079*** (0.015)	1.181*** (0.016)	1.195*** (0.016)	1.213*** (0.016)	1.203*** (0.016)
Fam income: 5-7 mw	1.806*** (0.020)		1.140*** (0.021)	1.266*** (0.022)	1.280*** (0.022)	1.302*** (0.022)	1.292*** (0.022)
Fam income: 7-10 mw	1.899*** (0.027)		1.122*** (0.028)	1.254*** (0.029)	1.282*** (0.029)	1.309*** (0.029)	1.299*** (0.029)
Fam income: >10 mw	1.380*** (0.026)		0.674*** (0.028)	0.912*** (0.034)	0.928*** (0.034)	0.932*** (0.035)	0.926*** (0.035)

**Tabla 21.** Coeficientes estimados de la regresión Logit. Variable dependiente: "¿Está en una IES?".

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Mother ed: secondary			0.456*** (0.006)	0.453*** (0.009)	0.447*** (0.009)	0.476*** (0.009)	0.475*** (0.009)
Mother ed: T&T			0.794*** (0.009)	0.759*** (0.013)	0.734*** (0.013)	0.779*** (0.013)	0.781*** (0.013)
Mother ed: University+			0.858*** (0.010)	0.820*** (0.015)	0.774*** (0.015)	0.840*** (0.015)	0.845*** (0.015)
Works (=1)			-0.229*** (0.010)	-0.246*** (0.016)	-0.259*** (0.016)	-0.266*** (0.016)	-0.266*** (0.016)
Female (=1)			-0.014*** (0.005)	0.013* (0.008)	0.009 (0.008)	0.008 (0.008)	0.007 (0.008)
Age			-0.188*** (0.002)	-0.132*** (0.003)	-0.132*** (0.003)	-0.132*** (0.003)	-0.132*** (0.003)
Lives in metropolitan area							0.120*** (0.014)

**Tabla 22.** Coeficientes estimados de la regresión Logit. Variable dependiente: “¿Está en una IES?”.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Sb11 & After	0.087*** (0.006)	-0.006 (0.006)	0.092*** (0.006)	0.047*** (0.006)	0.043*** (0.006)	0.053*** (0.006)	0.031*** (0.006)
Fam income: 1-2 mw & after	-0.094*** (0.014)		-0.147*** (0.014)	-0.225*** (0.014)	-0.222*** (0.015)	-0.211*** (0.015)	-0.193*** (0.015)
Fam income: 2-3 mw & after	-0.083*** (0.016)		-0.184*** (0.016)	-0.320*** (0.017)	-0.321*** (0.017)	-0.320*** (0.017)	-0.295*** (0.018)
Fam income: 3-5 mw & after	-0.192*** (0.019)		-0.262*** (0.020)	-0.457*** (0.021)	-0.462*** (0.021)	-0.467*** (0.022)	-0.458*** (0.023)
Fam income: 5-7 mw & after	-0.391*** (0.029)		-0.381*** (0.030)	-0.629*** (0.032)	-0.634*** (0.032)	-0.643*** (0.032)	-0.619*** (0.033)
Fam income: 7-10 mw & after	-0.641*** (0.040)		-0.488*** (0.041)	-0.762*** (0.043)	-0.779*** (0.043)	-0.787*** (0.043)	-0.777*** (0.045)
Fam income: >10 mw & after	-0.700*** (0.037)		-0.461*** (0.038)	-0.858*** (0.046)	-0.868*** (0.046)	-0.845*** (0.047)	-0.812*** (0.048)

**Tabla 23.** Coeficientes estimados de la regresión Logit. Variable dependiente: “¿Está en una IES?”.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Mother ed: secondary & after				0.006 (0.012)	0.028** (0.012)	-0.011 (0.012)	-0.013 (0.013)
Mother ed: T&T & after				0.082*** (0.018)	0.122*** (0.018)	0.058*** (0.018)	0.056*** (0.019)
Mother ed: University+ & after				0.069*** (0.020)	0.105*** (0.020)	0.010 (0.020)	-0.024 (0.021)
Works (=1) & after				0.017 (0.020)	0.012 (0.020)	0.027 (0.020)	0.037* (0.021)
Female (=1) & after				-0.043*** (0.010)	-0.037*** (0.010)	-0.034*** (0.010)	-0.025*** (0.011)
Age & after				-0.114*** (0.004)	-0.108*** (0.004)	-0.110*** (0.004)	-0.114*** (0.004)
Lives in MA & after							-0.152*** (0.018)
Department controls					X	X	X
Department controls & after						X	X
Observations	967,682	964,471	957,440	957,440	957,439	957,439	835,881

Analizando ahora la elección entre universidad oficiales y no oficiales (tablas 24 y 25), es posible darse cuenta que la probabilidad de ir a la universidad pública es mayor cuando los resultados en las pruebas Saber 11 son mejores e, igualmente, que entre más grande sean los ingresos del estudiante, menor es la probabilidad de ir a una institución educativa pública de educación superior. Otro resultado visible es que una madre menos educada aumenta la probabilidad de asistir a una universidad pública. Sin embargo, enfocándonos en las interacciones (tablas 26 y 27) para ver los cambios entre 2000 y 2009, se evidenció que la prueba Saber 11 se hizo menos importante para asistir a una institución pública, en razón de la expansión de la capacidad de las universidades y la reducción de la diferencia de entrada de los “mejores” entre instituciones públicas y privadas. Adicionalmente, el ingreso no es una variable relevante, excepto para las personas de un nivel de ingreso medio, quienes aumentaron su probabilidad de ingreso a la universidad pública.

**Tabla 24.** Coeficientes estimados de la regresión Logit. Variable dependiente: "Matrícula en una IES pública".

	(1)	(2)	(3)	all			(7)
				(4)	(5)	(6)	
Own sb11 score	0.455*** (0.007)	0.356*** (0.006)	0.473*** (0.007)	0.478*** (0.007)	0.584*** (0.008)	0.569*** (0.008)	0.571*** (0.008)
After (=1 cohort 2009)	-0.146*** (0.023)	0.097*** (0.021)	-0.188*** (0.024)	-0.018 (0.107)	0.545*** (0.112)	-0.124 (0.118)	0.078 (0.126)
Fam income: 1-2 mw	-0.500*** (0.023)		-0.366*** (0.023)	-0.382*** (0.023)	-0.269*** (0.024)	-0.280*** (0.025)	-0.273*** (0.025)
Fam income: 2-3 mw	-1.029*** (0.024)		-0.770*** (0.025)	-0.785*** (0.025)	-0.641*** (0.026)	-0.652*** (0.026)	-0.644*** (0.026)
Fam income: 3-5 mw	-1.543*** (0.025)		-1.180*** (0.027)	-1.190*** (0.028)	-1.093*** (0.029)	-1.090*** (0.029)	-1.082*** (0.029)
Fam income: 5-7 mw	-2.119*** (0.033)		-1.677*** (0.034)	-1.684*** (0.035)	-1.615*** (0.037)	-1.605*** (0.037)	-1.598*** (0.037)
Fam income: 7-10 mw	-2.515*** (0.042)		-1.942*** (0.043)	-1.941*** (0.045)	-1.936*** (0.048)	-1.924*** (0.048)	-1.917*** (0.048)
Fam income: >10 mw	-3.439*** (0.050)		-2.568*** (0.053)	-2.585*** (0.063)	-2.507*** (0.067)	-2.505*** (0.067)	-2.499*** (0.067)

**Tabla 25.** Coeficientes estimados de la regresión Logit. Variable dependiente: "Matrícula en una IES pública".

	(1)	(2)	(3)	all			(7)
				(4)	(5)	(6)	
Mother ed: secondary			-0.113*** (0.010)	-0.172*** (0.016)	-0.226*** (0.017)	-0.201*** (0.017)	-0.202*** (0.017)
Mother ed: T&T			-0.237*** (0.013)	-0.325*** (0.021)	-0.400*** (0.022)	-0.349*** (0.022)	-0.352*** (0.022)
Mother ed: University+			-0.321*** (0.013)	-0.424*** (0.023)	-0.596*** (0.024)	-0.536*** (0.024)	-0.542*** (0.025)
Works (=1)			-0.032* (0.017)	0.050* (0.030)	0.077** (0.032)	0.062** (0.032)	0.063** (0.032)
Female (=1)			-0.259*** (0.007)	-0.161*** (0.013)	-0.174*** (0.014)	-0.181*** (0.014)	-0.181*** (0.014)
Age			0.011*** (0.003)	0.010** (0.004)	0.036*** (0.005)	0.022*** (0.005)	0.023*** (0.005)
Lives in metropolitan area							-0.121*** (0.026)

**Tabla 26.** Coeficientes estimados de la regresión Logit. Variable dependiente: "Matrícula en una IES pública".

	(1)	(2)	(3)	all			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Sb11 & After	-0.118*** (0.008)	-0.042*** (0.008)	-0.086*** (0.008)	-0.093*** (0.009)	-0.110*** (0.009)	-0.088*** (0.010)	-0.082*** (0.010)
Fam income: 1-2 mw & after	-0.023 (0.026)		0.019 (0.027)	0.040 (0.027)	0.006 (0.028)	0.026 (0.029)	0.029 (0.029)
Fam income: 2-3 mw & after	0.035 (0.028)		0.121*** (0.029)	0.142*** (0.030)	0.126*** (0.031)	0.145*** (0.032)	0.143*** (0.033)
Fam income: 3-5 mw & after	0.082*** (0.031)		0.244*** (0.032)	0.257*** (0.034)	0.322*** (0.036)	0.327*** (0.036)	0.328*** (0.037)
Fam income: 5-7 mw & after	0.019 (0.043)		0.302*** (0.044)	0.311*** (0.046)	0.458*** (0.049)	0.457*** (0.050)	0.456*** (0.051)
Fam income: 7-10 mw & after	-0.065 (0.058)		0.237*** (0.059)	0.241*** (0.063)	0.541*** (0.067)	0.543*** (0.067)	0.549*** (0.069)
Fam income: >10 mw & after	-0.161** (0.073)		0.100 (0.075)	0.123 (0.086)	0.440*** (0.091)	0.448*** (0.091)	0.398*** (0.094)

**Tabla 27.** Coeficientes estimados de la regresión Logit. Variable dependiente: "Matrícula en una IES pública".

	(1)	(2)	(3)	all			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Mother ed: secondary & after				0.098*** (0.020)	0.137*** (0.021)	0.105*** (0.022)	0.109*** (0.023)
Mother ed: T&T & after				0.144*** (0.027)	0.225*** (0.028)	0.167*** (0.028)	0.178*** (0.030)
Mother ed: University+ & after				0.165*** (0.028)	0.212*** (0.030)	0.137*** (0.031)	0.142*** (0.032)
Works (=1) & after				-0.121*** (0.036)	-0.053 (0.038)	-0.035 (0.038)	-0.043 (0.040)
Female (=1) & after				-0.152*** (0.016)	-0.172*** (0.017)	-0.165*** (0.017)	-0.162*** (0.017)
Age & after				0.001 (0.006)	-0.027*** (0.006)	-0.011* (0.006)	-0.013** (0.007)
Lives in MA & after							-0.158*** (0.031)
Department controls					X	X	X
Department controls & after						X	X
<i>Observations</i>	340,200	339,542	337,348	337,348	337,348	337,348	299,595
<i>pseudo-R-squared</i>	0.079	0.063	0.094	0.095	0.169	0.172	0.174

Finalmente, se analizaron las diferencias en la distribución de aquellos que van a programas nuevos versus los que van a programas existentes (tabla 28 y 29). Los resultados muestran que los programas nuevos están atrayendo personas con un menor Saber 11 y de escasos recursos. Otro resultado relevante señalado por Carranza hace referencia a que, en estos diez años, vivir en un área metropolitana perdió importancia.

**Tabla 28.** Coeficientes estimados de la regresión Logit. Programas de 2 y 4 años.  
Variable dependiente: "Matrícula en un programa nuevo".

	all					
	(1)	(2)	(3)	(4)	(5)	(6)
Own sb11 score	-0.299*** (0.005)	-0.308*** (0.005)	-0.274*** (0.006)	-0.285*** (0.006)	-0.155*** (0.007)	-0.137*** (0.008)
Fam income: 1-2 mw	-0.059*** (0.014)		0.006 (0.015)	-0.021 (0.017)	0.005 (0.019)	0.014 (0.020)
Fam income: 2-3 mw	-0.195*** (0.016)		-0.058*** (0.018)	-0.116*** (0.020)	-0.062*** (0.023)	-0.052** (0.023)
Fam income: 3-5 mw	-0.383*** (0.020)		-0.173*** (0.022)	-0.241*** (0.025)	-0.162*** (0.028)	-0.152*** (0.028)
Fam income: 5-7 mw	-0.638*** (0.030)		-0.358*** (0.034)	-0.463*** (0.037)	-0.374*** (0.040)	-0.357*** (0.040)
Fam income: 7-10 mw	-0.775*** (0.043)		-0.460*** (0.047)	-0.609*** (0.051)	-0.533*** (0.053)	-0.506*** (0.053)
Fam income: >10 mw	-1.000*** (0.045)		-0.650*** (0.052)	-0.791*** (0.057)	-0.757*** (0.059)	-0.738*** (0.060)

**Tabla 29.** Coeficientes estimados de la regresión Logit. Programas de 2 y 4 años.  
Variable dependiente: "Matrícula en un programa nuevo".

	all					
	(1)	(2)	(3)	(4)	(5)	(6)
Works (=1)			0.073*** (0.021)	0.065*** (0.025)	0.014 (0.028)	0.015 (0.028)
Female (=1)			0.063*** (0.010)	0.056*** (0.012)	0.033** (0.014)	0.029** (0.014)
Age			0.021*** (0.004)	0.022*** (0.005)	0.020*** (0.005)	0.017*** (0.006)
Lives in metropolitan area				-0.062*** (0.019)	-0.182*** (0.022)	-0.107*** (0.023)
HEI: T&T					1.229*** (0.025)	1.256*** (0.026)
HEI: University/Technical					1.051*** (0.017)	1.083*** (0.018)
HEI: Public					-0.526*** (0.017)	-0.588*** (0.018)

Para concluir, se busca obtener una descomposición de los choques de oferta y demanda, donde se denomina oferta de educación superior a la probabilidad de que un estudiante con ciertas características se matricule en un programa. Podemos decir que su demanda aumentó si las características del estudiante mejoran y su probabilidad de matricularse aumenta. Se realizó una estimación de la probabilidad de matrícula en cualquier tipo de programa para las cohortes de 2000 y 2009. Inferimos los cambios en la demanda usando las probabilidades de 2000 con los estudiantes de 2009. Un resultado importante es que la probabilidad de asistir a la universidad en 2000 para una persona de escasos recursos con un nivel alto era de 25%, mientras que en 2009 es 45%. Al hacer una descomposición del crecimiento en las matrículas se puede evidenciar que el cambio más grande se da cuando entran 8.386 estudiantes más en 2009 respecto al año 2000 y de estos, los que se pueden atribuir al cambio en las características de los estudiantes son 3.904, aquellos que de todas formas habrían asistido. Los beneficiarios son los 4.482 estudiantes que pudieron asistir gracias a la ampliación de la cobertura (tabla 30 y 31).

**Tabla 30.** Probabilidades predichas de matricularse a una IES.

Probabilidades predichas de matricularse a una IES					
Cohort 2000, T=2000					
	ability 1	ability 2	ability 3	ability 4	ability 5
5+ smlmv	0.313	0.393	0.463	0.549	0.638
3-5 smlv	0.251	0.32	0.392	0.479	0.61
2-3 smlv	0.173	0.229	0.287	0.364	0.504
1-2 smlv	0.109	0.148	0.19	0.252	0.385
<1 smlv	0.058	0.081	0.107	0.149	0.258
Cohort 2009, T=2009					
	ability 1	ability 2	ability 3	ability 4	ability 5
5+ smlmv	0.344	0.462	0.562	0.658	0.727
3-5 smlv	0.358	0.475	0.569	0.669	0.769
2-3 smlv	0.292	0.394	0.479	0.585	0.705
1-2 smlv	0.197	0.275	0.349	0.451	0.591
<1 smlv	0.124	0.177	0.231	0.315	0.459

**Tabla 31.** Cambios en la cantidad de matrículas.

Cambios en la cantidad de matrículas					
Baseline change in enrollment between 2000-2009					
	ability 1	ability 2	ability 3	ability 4	ability 5
5+ smlmv	-232	-239	-138	227	1,954
3-5 smlv	45	306	670	1,540	3,366
2-3 smlv	683	1,417	2,219	3,687	5,095
1-2 smlv	2,867	4,446	6,047	8,055	8,386
<1 smlv	1,676	2,447	2,981	3,525	3,035
Change attributed to the change in demand					
	ability 1	ability 2	ability 3	ability 4	ability 5
5+ smlmv	-269	-371	-455	-595	-1,060
3-5 smlv	-133	-114	-51	251	1,140
2-3 smlv	130	317	563	1,132	2,189
1-2 smlv	1,256	1,692	2,256	2,966	3,904
<1 smlv	640	893	1,058	1,213	1,308

### *Valor agregado en la educación superior*

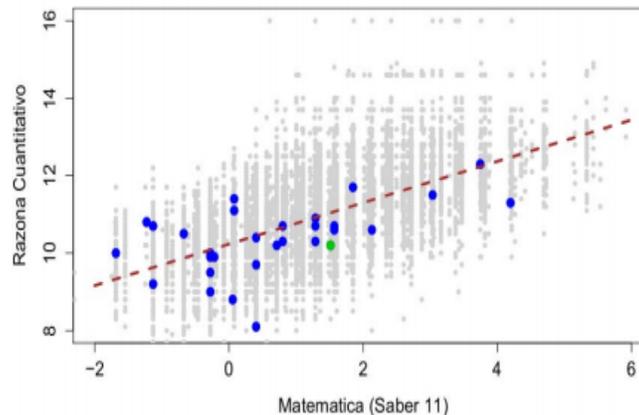
Ernesto San Martín, Pontificia Universidad Católica de Chile

La idea de la presentación es cómo comunicar indicadores. En relación con el valor agregado, el conferencista señala inicialmente alguno de los usos de las medidas del valor agregado: (i) analizar la efectividad de sistemas educacionales, (ii) relacionar la efectividad de las escuelas con aspectos específicos de un sistema educacional, (iii) evaluar sistemas de responsabilidad y (iv) evaluar críticamente aspectos políticos educacionales. En el caso colombiano, el decreto 3969 da algunas indicaciones generales de por qué se quisieran emplear medidas de valor agregado, en razón de que pueden resultar útiles para generar indicadores de calidad de la educación de programas e instituciones de educación superior, y así poder comparar las instituciones. Las medidas de crecimiento permiten generar información que apruebe la comparación entre metodologías de enseñanza y la cualificación de los procesos institucionales como la formulación de políticas, así como asegurar que dichas comparaciones puedan realizarse en el transcurso del tiempo.

Los modelos de valor agregado típicamente se asocian a la rendición de cuentas y a una percepción de desconfianza del uso de estas herramientas estadísticas para monitorear la calidad de la educación por la presión que pueden generar en las instituciones. Así, aparecen ciertas críticas a las medidas del valor agregado, como la presión externa en detrimento de la enseñanza de la institución de educativa y la omisión de temas relevantes producto del énfasis en las áreas del sistema de monitoreo, que, según San Martín, podrían ser problemas de comunicación de los fenómenos que están sucediendo.

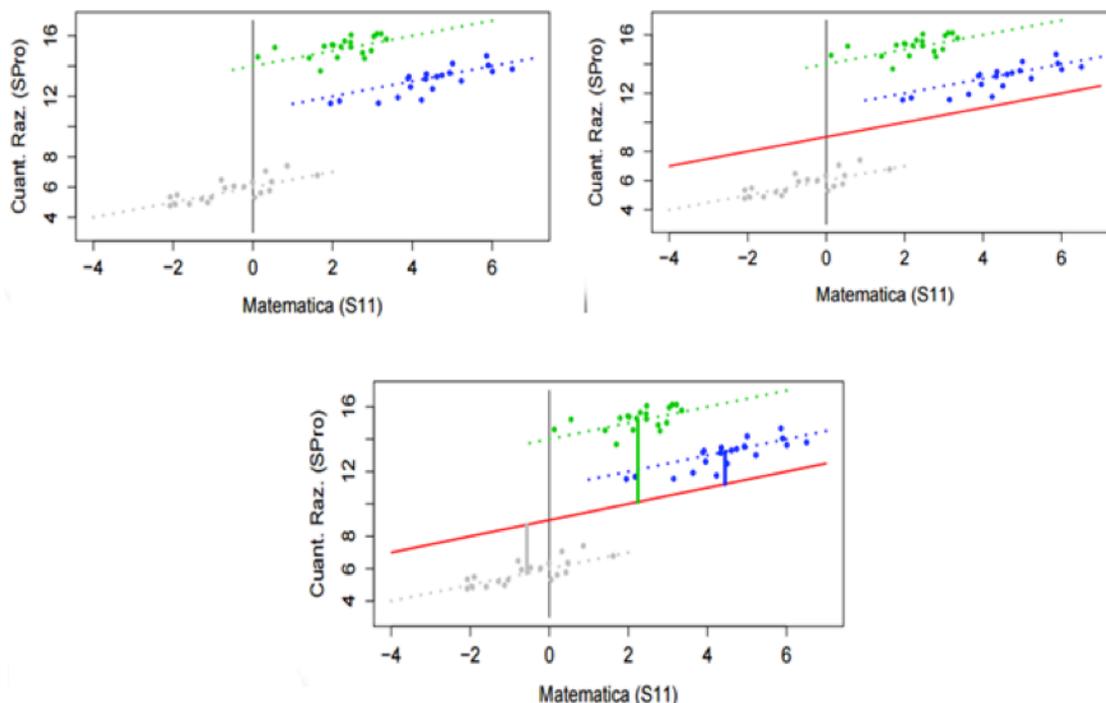
Colombia es de los pocos países que miden a los estudiantes al culminar su educación universitaria, la prueba Saber Pro. En esta se evalúa a los estudiantes en cinco niveles de competencia: comunicación escrita, razonamiento cuantitativo, inglés, lectura crítica y competencias ciudadanas. Igualmente el conferencista resalta la importancia del proceso de evaluación mediante Saber 11 antes de entrar a la educación terciaria.

Para comenzar a ilustrar el tema central de la presentación, muestra un ejemplo para la carrera de ingeniería (figura 13), en la que el eje vertical mide razonamiento cuantitativo y matemáticas en el eje horizontal. San Martín señala que hay escalas distintas pero lo importante es el patrón. En este caso, el punto verde es un estudiante con ciertas características: reside en Bogotá, asiste a una universidad X, proviene del colegio Y, es de estrato socio-económico 5, tiene un nivel sociocultural (Inse) 80.15, su padre es jubilado y su madre trabaja en un nivel directivo. El estudiante tiene asociado un puntaje de razonamiento cuantitativo y de matemáticas de Saber 11. Los puntos azules representan los compañeros del estudiante en la universidad X.

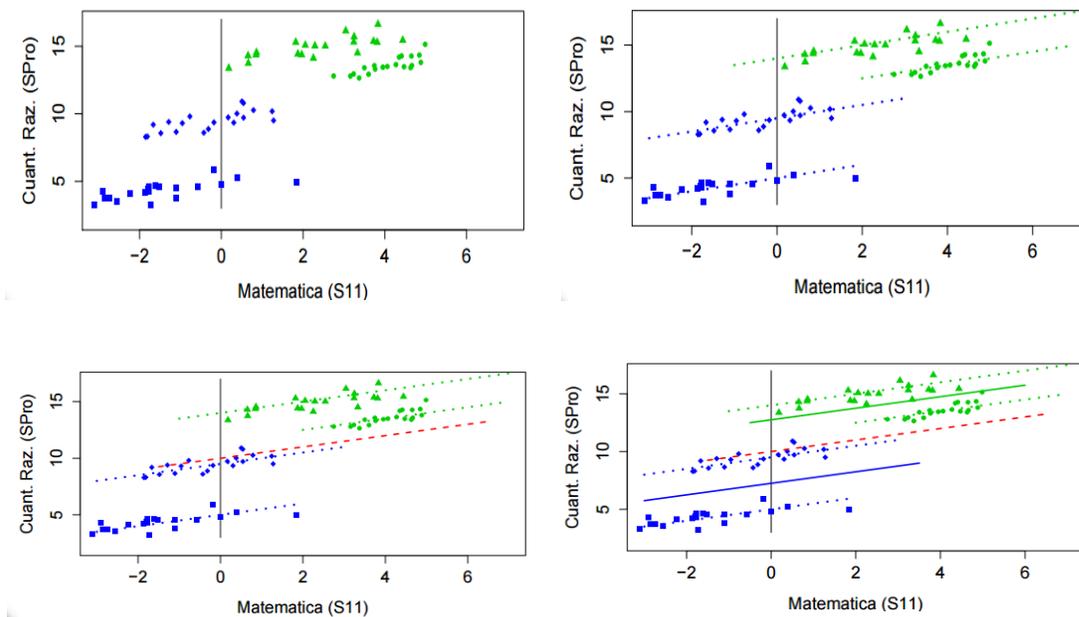
**Figura 13.** Valor agregado “¿Cuál es el resultado esperado?”.

El paso siguiente es entender el concepto de “valor agregado” como la comparación entre el resultado esperado de Saber Pro en una universidad específica con el resultado esperado de Saber Pro en la universidad promedio. La manera intuitiva de entenderlo es de forma gráfica, donde el eje vertical de la figura 14a se tiene razonamiento cuantitativo y matemáticas en el eje horizontal. Cada color representa una universidad diferente, la universidad gris, cuyo rendimiento en razonamiento cuantitativo no es muy bueno; la universidad verde, cuyos resultados son muy buenos en Saber Pro; y la universidad azul, un poquito más abajo que las universidades verdes pero ambas más arriba que la universidad gris. Lo que podemos observar con valor agregado es la relación entre matemáticas Saber Pro y razonamiento cuantitativo en la misma prueba para cada una de las universidades, que son las líneas punteadas que aparecen en la figura 14a.

Luego se hace una abstracción del hecho de que cada estudiante pertenece a una universidad, vemos los estudiantes como un todo y dibujamos la relación que existe entre matemáticas Saber 11 y razonamiento cuantitativo en Saber Pro, que es la línea roja (figura 14b). En consecuencia el valor agregado es la comparación entre la línea gris y la línea roja, y podría interpretarse que respecto a la línea roja, la universidad gris no agrega. No obstante si nos fijamos en la universidad azul, podríamos decir que respecto a la línea roja la universidad azul agrega algo, pero siempre con un punto de referencia (figura 14c).

**Figura 14.** Valor agregado “¿Cuál es el resultado esperado?”.

San Martín pone otro ejemplo señalando el uso de la variable Saber 11 (figura 15). En este caso se tienen cuatro universidades distintas, representadas por triángulos, círculos, rombos y cuadrados, que tienen características en común: la universidad de los cuadrados y la de los rombos tienen un promedio de matemáticas en Saber 11 muy similar, así como las universidades representadas con triángulos y círculos. Si la línea de referencia vuelve a ser la roja, observamos que, por muy poco, la universidad de los rombos no agrega. Pero si además de tener en cuenta el puntaje Saber 11 consideramos el hecho que su promedio en matemáticas es similar, podemos dibujar la línea azul continua, que es la nueva referencia para cuadrados y rombos, mientras que la línea verde continua es la nueva referencia para triángulos y círculos. Ahora, en comparación a la situación anterior, la universidad de los rombos agrega valor cuando en la situación con la línea roja no lo hacía; ilustrando que las medidas de crecimiento dependen del contexto contra el cual se desean comparar y es definido por la política pública.

**Figura 15.** El papel de las covariables y su impacto en el VA.

En los análisis realizados por el Icfes se han hablado de dos modelos: (i) el modelo 1, en el que se ha escogido una determinada batería de variables que definen la referencia usando solamente Saber 11; y (ii) el modelo 2, en el que, además de las medidas individuales de Saber 11, se compone por lo que se denomina un efecto composicional, un promedio del Saber Pro de los estudiantes que pertenecen a una determinada carrera universitaria. El modelo 1 proporciona el valor agregado de estar en una institución educativa, aunque dicho valor agregado no se debe a la calidad de la formación en sí, mientras que el modelo 2 proporciona el valor agregado de estar en una institución educativa, en la que dicho valor agregado no refleja los mecanismos de selección de la institución sino lo que ocurre después de la selección.

Para comunicar esta información puede considerarse los vecinos cercanos, idea que consiste en que una universidad o programa universitario se compare con vecinos escogidos por información socioeconómica, o mostrar la información proporcionada por ambos modelos de forma compacta y simultánea. El conferencista puso como ejemplo el programa de derecho y como variable dependiente la lectura crítica (figura 16). Los círculos son los vecinos cercanos con los que esa determinada universidad se compara con otras en indicadores socioeconómicos. En el panel de la izquierda tenemos el modelo 1 y en el de la derecha el modelo 2. Se puede afirmar que esta universidad agrega más valor respecto a las otras en el panel de la izquierda, cuando las medidas de crecimiento son un indicativo para los padres que están tomando la decisión de

escoger una institución educativa porque solo se está controlando por el nivel inicial de los estudiantes. En el panel de la derecha cambia un poco el patrón, se controla por lo que el grupo puede contribuir y se observa que institución está de tercera o cuarta en el ranking.

Pero, ¿cómo se puede dar información para que la universidad se mire hacia dentro y trate de mejorar para proveer una mejor calidad de la educación para sus estudiantes? San Martín propone un gráfico (figura 17) en donde el eje horizontal es el modelo 1 y el eje vertical es el modelo 2 con cuatro cuadrantes: (i) en el primero se es efectivo en el modelo 1 y en el modelo 2; (ii) en el segundo se es inefectivo cuando se trata de los padres (modelo 1), pero efectivo en relación a los términos pedagógicos (modelo 2); (iii) en el tercer cuadrante se es inefectivo en ambos modelos; y (iv) en el cuarto es efectivo el modelo de selección por parte de padres e inefectivo en el modelo pedagógico. La línea punteada sería la mejor universidad o programa en sus vecinos, pero el problema es que hay una arbitrariedad en cuanto a cuál competencia se escoge, ya que hay cuatro variables posibles de competencia, lo que nos lleva a preguntarnos si existe alguna forma de usar toda la información. En razón a que todos los estudiantes son evaluados en cinco competencias, es posible usar la información disponible en un modelo de aporte relativo multidimensional. Este modelo se comunica a través de una pirámide (figuras 18 y 19), en la que cada línea representa una competencia: si es una línea verde, aporta en esa dirección; si es una línea roja, no agrega.

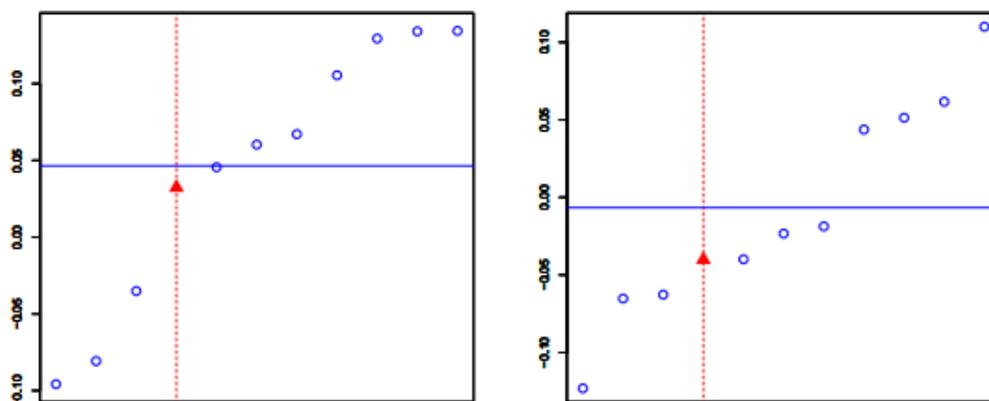
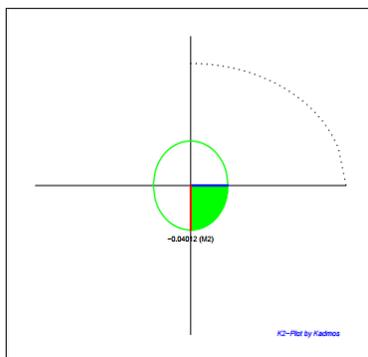
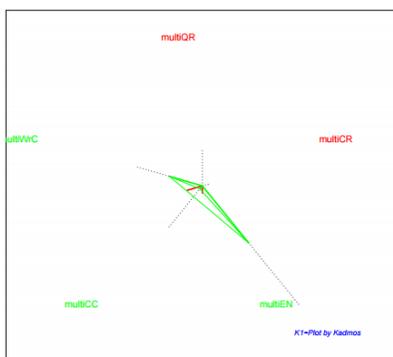


Figura 16. Modelos 1 y 2 (derecho y lectura crítica).

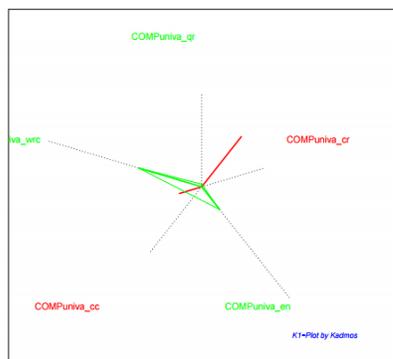
**Figura 17.** Kadmos Plot © para modelos 1 y 2.



**Figura 18.** Kadmos Plot © para el modelo 1.



**Figura 19.** Kadmos Plot © para el modelo 2.



## Conferencias sobre índices de calidad

*Aprendizaje de monitoreo para todos. Retos y oportunidades de los índices compuestos que sirven a los objetivos de desarrollo sostenible*

Silvia Montoya, Unesco

La presentación se centra en exponer el entorno político, explicar la necesidad del uso de índices compuestos e ilustrar la experiencia argentina en el uso de índices de calidad, más específicamente a través de la Unidad de la Evaluación Integral de Calidad y Equidad Educativa (IECEP). Inicialmente, la agenda de 2015 reconoce explícitamente la pobreza, la desigualdad, la seguridad alimentaria, un medio ambiente sostenible y los derechos humanos como universales. En este contexto, para generar un desarrollo sostenible, se evidencia la necesidad de datos sostenibles; además de observar el requerimiento de más datos a nivel local para ayudar al planeamiento estratégico, la investigación y la rendición de cuentas. Lo anterior plantea un paradigma de cambio, pues las estadísticas nacionales sobre el desarrollo se deben basar en desagregaciones de datos subnacionales.

En esta era de grandes bases de datos, Montoya señala algunas complejidades que se presentan en su manejo como: (i) datos que sirven a política pública pueden ser inconsistentes, inefectivos y no tener interpretación; (ii) la información puede ser obesa, es decir, es abundante, dispersa y de mala calidad; y (iii) los hacedores de política no hacen uso de los datos. Montoya hace un repaso del proceso de identificación de indicadores globales para ODS, en la Comisión Estadística de la ONU, como se describe en la figura 20 y que responde a unos niveles de monitoreo propuesto.

**Figura 20.** Proceso de identificación de indicadores globales para ODS en la Comisión Estadística de la ONU.



La funcionalidad de los indicadores de desarrollo sostenible es supervisar el progreso mundial alcanzado sobre los 17 nuevos objetivos para el desarrollo sostenible y las 169 metas que buscan erradicar la pobreza, combatir las desigualdades y promover la prosperidad, al tiempo que protegen el medio ambiente de aquí a 2030. En relación a los temas de calidad y equidad reflejados en los indicadores globales, Montoya señala respecto al primero que deberíamos enfocarnos en un grupo pequeño de medidas de resultados, tener medidas directas de aprendizaje y habilidad, y considerar otras medidas de calidad; mientras para la equidad propone que los indicadores deben ser desagregados cuando sea posible, fortaleciendo los mensajes en la igualdad de género, y deben asegurar que la desventaja múltiple sea capturada y que los grupos desfavorecidos y contextos difíciles de medir reciban una atención especial.

Montoya muestra algunos índices usados en otros países diferentes de Argentina, como el progreso anual adecuado creado luego del ley de 2001 "Que ningún niño se quede atrás" y que varía según el estado en Estados Unidos, el índice de desarrollo de la educación básica (IDEB) en Brasil y el índice de mejoramiento de la escuela secundaria (IMESA) en Argentina, el cual incluye variables como el tiempo promedio en culminar los estudios de secundaria, la tasa de graduación y los puntajes de la evaluación nacional Operativo Nacional de Evaluación (ONE).

En Buenos Aires, con el objetivo de monitorear mejoras en las escuelas, las comunas y los distritos escolares, además de poder establecer metas propias, se crea el índice de equidad y calidad de la educación porteña (IECEP). El IECEP es un índice desarrollado por la Dirección General de Evaluación de la Calidad Educativa que permite obtener una medida multidimensional del desempeño de las escuelas. Esta característica del IECEP es fundamental ya que al ser la calidad de la educación una variable multidimensional, se necesita un instrumento que permita abarcar y medir diferentes aspectos de la educación.

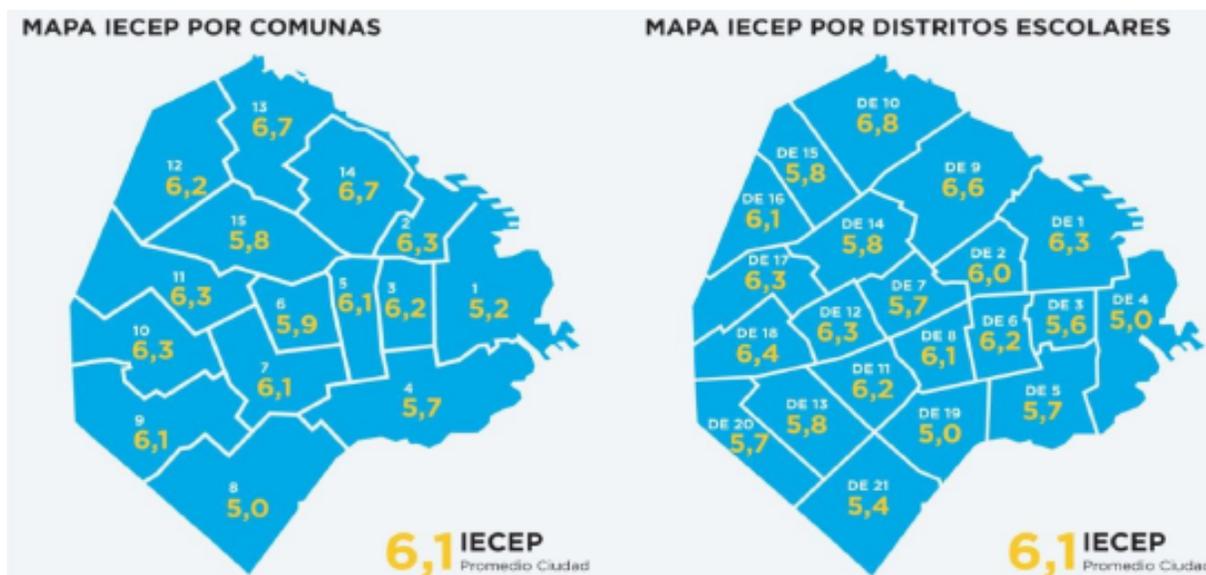
El IECEP justifica su utilidad porque es una medida fácil y simple de interpretar la realidad. Valores altos del índice sólo se logran teniendo simultáneamente buenas trayectorias de los alumnos, buenos resultados de aprendizaje y poca diferencia en relación con el promedio de la ciudad. Hay tres dimensiones que componen el índice: logros de aprendizaje (resultados en lengua y matemática del ONE), la trayectoria de los alumnos (tasas de promoción y repetición de los alumnos, datos que completan las escuelas en el relevamiento anual) y la equidad del aprendizaje (distancia entre los resultados de una escuela y los del promedio de la ciudad en las evaluaciones del ONE).

Adicionalmente, Montoya explica que el IECEP resulta de la sumatoria de ocho indicadores cuyos valores ponderados suman 10. Los indicadores seleccionados son: los porcentajes de alumnos que obtuvieron nivel alto y medio en las evaluaciones ONE de lengua y matemática, la diferencia entre estas dos, la tasas de promoción y repetición, y las diferencias entre los resultados de la escuela y los de la ciudad en las pruebas ONE de lengua y matemáticas en porcentajes. Observando la figura 21 se puede evidenciar que el promedio de la ciudad es 6,1 y

que las comunas 13 y 14 presentan el mejor promedio en contraste a la comuna 8 y 1, que presentan el IECEP más débil.

Otros resultados obtenidos a través del cálculo de este índice hacen referencia a que muchas escuelas de gestión estatal se encuentran en una mejor situación que las de gestión privada y que, en relación con el nivel socioeconómico de la familia, hay una relación positiva y significativa que influye sobre el capital humano inicial, y sobre el que los estudiantes continúan incorporando en sus hogares a lo largo de su vida estudiantil. Finalmente, Montoya concluye su ponencia señalando algunas estrategias relevantes para mejorar el aprendizaje, como la construcción de consenso: es clave la participación de los interesados en el tema –directores, profesores, sindicatos, sociedad civil y hacedores de política–, y el desarrollo de técnicas de liderazgo, como por ejemplo, la aplicación de herramientas comunes, metodologías, marcos.

**Figura 21.** IECEP de la ciudad de Buenos Aires.



### *El IDEB y educación nacional. Retos y perspectivas*

Alexandre dos Santos, director de Evaluación y Educación Básica del INEP, Brasil

Los temas abordados en esta ponencia son las características del índice de desarrollo de la educación básica (IDEB), el mapa de la calidad de la educación, la evaluación en el Plan Nacional de Educación y los desafíos y perspectivas. El IDEB es la combinación de dos indicadores: (i) el desempeño de los alumnos en las pruebas estandarizadas (Prueba Brasil y SAEB) y (ii) la tasa de aprobación de los alumnos.

A mayor desempeño de los alumnos en las pruebas y mayor tasa de aprobación, más elevado será el IDEB, y para mejorar este indicador es necesario que las redes de enseñanza y las escuelas trabajen simultáneamente en las dos dimensiones. El IDEB tiene un rango de 0 a 10.

Técnicamente, el IDEB puede representarse de la siguiente forma:  $Ideb_{ij} = N_{ij} * P_{ij}$ , donde  $Ideb_{ij}$  es el índice de desarrollo de la educación básica de la etapa/nivel de enseñanza  $i$  en la unidad  $j$  (donde  $j$  puede ser escuela, red de enseñanza, región geográfica, unidad de la federación, el país, etc.).  $N_{ij}$  es el promedio estandarizado de competencia (Prueba Brasil o SAEB) de los alumnos en la etapa/nivel de enseñanza  $i$  en la unidad, y se calcula como:

$$N_{ij} = \frac{Rendimiento_{ij} - Min}{Max - Min} * 10$$

$P_{ij}$  es la tasa media de aprobación en la etapa/nivel de enseñanza  $i$  en la unidad  $j$ . Dos Santos muestra tres ejemplos hipotéticos (tabla 32), en los cuales, siguiendo el proceso anterior, se puede calcular el IDEB.

Las metas definidas apuntan a un camino trazado de evolución individual de los índices para que Brasil alcance niveles educativos similares a los que ahora tienen los países de la OCDE. En términos numéricos, eso significa evolucionar de una media nacional 4,2, registrada en 2007, a un IDEB de 6,0 en la primera fase de la educación fundamental en 2021. Brasil tiene metas claras para identificar si el país está mejorando y avanzando en calidad. Estos objetivos fueron establecidos en 2005 y son evaluados cada dos años por el IDEB. Para comprobar si se alcanzará el objetivo de 2021, las metas no se establecieron solamente a nivel país sino que se llevaron a nivel de escuela, municipio y estado.

El Instituto Nacional de Estudios e Investigaciones Educativas Anísio Teixeira también estableció objetivos intermedios para ser alcanzados por Brasil, los cuales se basan en tres premisas: (i) las trayectorias del IDEB en Brasil y en todas las redes tienen el comportamiento de una función logística, (ii) las trayectorias del IDEB por sistema escolar deben contribuir a reducir las desigualdades en términos de la calidad educativa y (iii) para que Brasil alcance el objetivo fijado, el esfuerzo de cada sistema escolar, estatal o local debe contribuir a objetivos individuales diferenciados. En la tabla 33 y figura 22 se pueden observar las proyecciones del IDEB para Brasil.

Dos Santos muestra que entre el período comprendido entre 2005 y 2013, a nivel municipal, ha existido un incremento de la calidad significativo evidenciado en el mayor número de municipios que tienen un indicador por encima de 5 en 2013, en comparación con años como 2007 y 2005, donde el panorama era bastante negativo. Finalmente, el ponente concluye exponiendo los principales desafíos y perspectivas cuestionando cómo debe ser la administración del Sistema Nacional de Evaluación de la Educación Básica, cómo construir una red de colaboración en evaluación con los estados y municipios, y cómo construir procesos de evaluación institucional y contextual.

**Tabla 32.** Cálculo del IDEB.

Ejemplo 1	Escuela A	Escuela B	Escuela C
Desempeño Medio Estandarizado	4,8	4,8	4,8
Tasa Media De Aprobación	0,9	0,8	0,5
<b>Ideb</b>	<b>4,3</b>	<b>3,8</b>	<b>2,4</b>

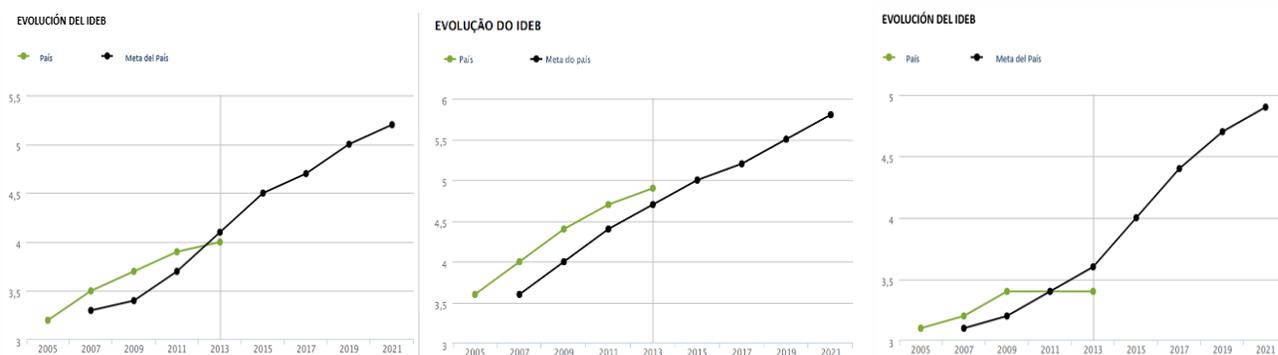
Ejemplo 2	Escuela A	Escuela B	Escuela C
Desempeño Medio Estandarizado	6,0	5,0	4,0
Tasa Media De Aprobación	0,7	0,7	0,7
<b>Ideb</b>	<b>4,2</b>	<b>3,5</b>	<b>2,8</b>

Ejemplo 3	Escuela A	Escuela B	Escuela C
Desempeño Medio Estandarizado	6,0	5,5	4,2
Tasa Media De Aprobación	1,0	1,0	1,0
<b>Ideb</b>	<b>6,0</b>	<b>5,5</b>	<b>4,2</b>

**Tabla 33.** Metas del IDEB por fase de enseñanza. Fuente: QEDU Academia.

Fase de enseñanza	IDEB 2005	Meta IDEB Brasil (red pública y privada)	Año de alcance de la meta
Años iniciales de la enseñanza fundamental	3.8	6.0	2021
Años finales de la enseñanza fundamental	3.5	6.0	2021
Enseñanza media	3.4	6.0	2021

**Figura 22.** Proyecciones IDEB 2021.

### *Índice sintético de calidad educativa, ISCE*

Silvana Godoy y Jorge Duarte, Subdirección de Análisis y Divulgación del Icfes

El índice sintético de calidad educativa (ISCE) nace del deseo de obtener una medida sencilla de la calidad de la educación, con el propósito de medir la asistencia a la escuela y el aprendizaje. En este proceso se trabajó observando las experiencias nacionales y de otros países, e igualmente se hizo uso de los datos disponibles más confiables. El resultado fue el diseño de tres índices para cada colegio en cada ciclo escolar –primaria, secundaria y media–, a nivel de institución educativa, municipio, etc. En la figura 23 se resumen las características más relevantes del ISCE.

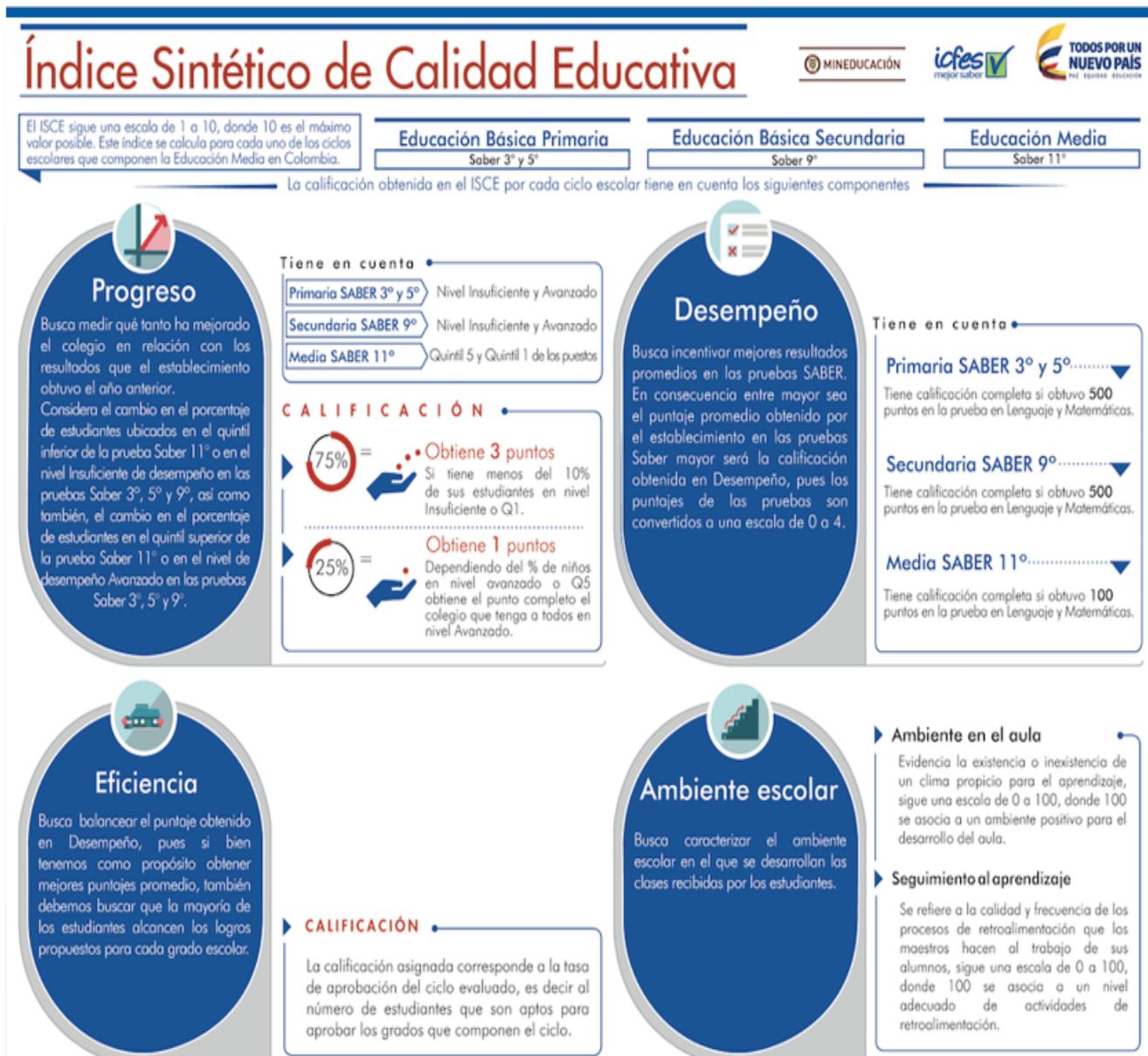
En relación con el ambiente escolar, Godoy hace hincapié en el seguimiento al aprendizaje de los estudiantes. Este mide la frecuencia con la que los estudiantes reportan que el docente realiza diferentes actividades de retroalimentación, las cuales permiten el desarrollo de la responsabilidad y el reforzamiento de los aprendizajes adquiridos en las clases. Un ejemplo que se observa en el área de lenguaje es cuando se les pregunta a los alumnos de grado quinto: “¿Cada cuánto tu profesor(a) de Lenguaje o Español realiza las siguientes actividades con las tareas que les deja para la casa?”. En la figura 24 se puede observar el tipo de respuesta dada por los niños y la frecuencia por clase en la que se dan dichas situaciones.

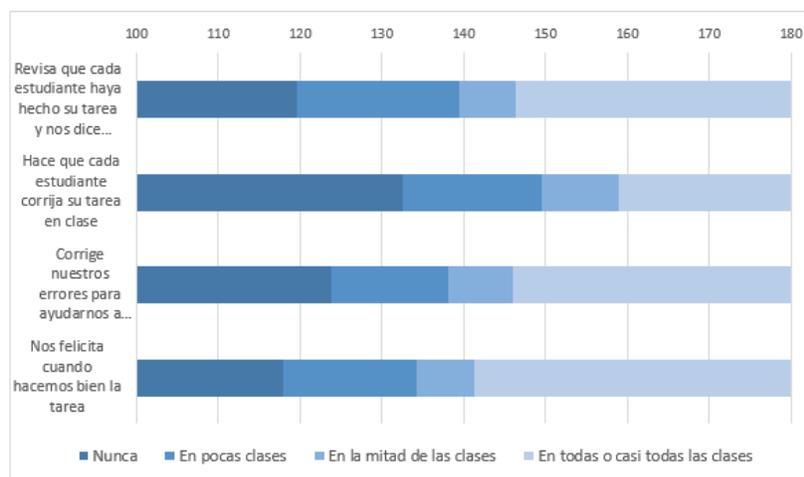
De igual forma, en relación con el ambiente en el aula, se da cuenta de la frecuencia con la que los estudiantes reportan la presencia de situaciones en las que se enmarca el desarrollo de las actividades de enseñanza y aprendizaje.

Un tema adicional explicado por Godoy es la metodología para el mejoramiento mínimo anual (MMA), cuyo objetivo es incentivar el reporte de información por parte de los colegios y, al mismo tiempo, imponer metas moderadas que implican un esfuerzo académico real. Se señaló que inicialmente, para colegios con datos faltantes, se imputó la media nacional en los componentes de eficiencia y ambiente escolar, además de imponer un cuarto de la meta de 2017 en la de 2016. En segundo lugar, se realizó un proceso de equivalencia respecto de Chile y se calculó el desempeño y, dada la elasticidad del puntaje respecto de los niveles de desempeño, el progreso. Para eficiencia y ambiente escolar, la meta se fijó manteniendo el promedio actual nacional a lo largo del tiempo. El cuadro muestra las metas impuestas para el año 2025 (tabla 34).

Finalmente, se presentaron algunas cifras generales y una clasificación de las secretarías con los ISCES más altos (Mosquera, Sabaneta, Chía, Duitama, Envigado, Floridablanca, Bucaramanga, Pasto, Rionegro y Zipaquirá) y los más bajos (Guainía, Magangué, Maicao, Buenaventura, Vichada, Tumaco, Uribe, Amazonas, Vaupés). Igualmente se diagnosticaron colegios que enfrentan retos en educación, identificando, por ejemplo, que 465 de 18.592 colegios tienen el mínimo puntaje en básica primaria, 365 de 10.770 tienen el mínimo puntaje en básica secundaria y 305 de 9.919 tienen un puntaje de 2 en media.

Figura 23. Índice sintético de calidad educativa.



**Figura 24.** Ejemplo de seguimiento al aprendizaje de los estudiantes.**Tabla 34.** Proyecciones de meta nacional 2015.

<b>Meta Completa nacional para el año 2025</b>			
	<b>Primaria</b>	<b>Secundaria</b>	<b>Media</b>
<b>Progreso</b>	2,588	2,910	2,525
<b>Desempeño</b>	2,785	2,920	2,754
<b>Ambiente Escolar</b>	0,758	0,757	
<b>Eficiencia</b>	0,876	0,827	1,777
<b>Total</b>	7,008	7,414	7,056

*La experiencia de Medellín en la medición de la calidad educativa. El índice multidimensional de calidad educativa como herramienta para la toma de decisiones*

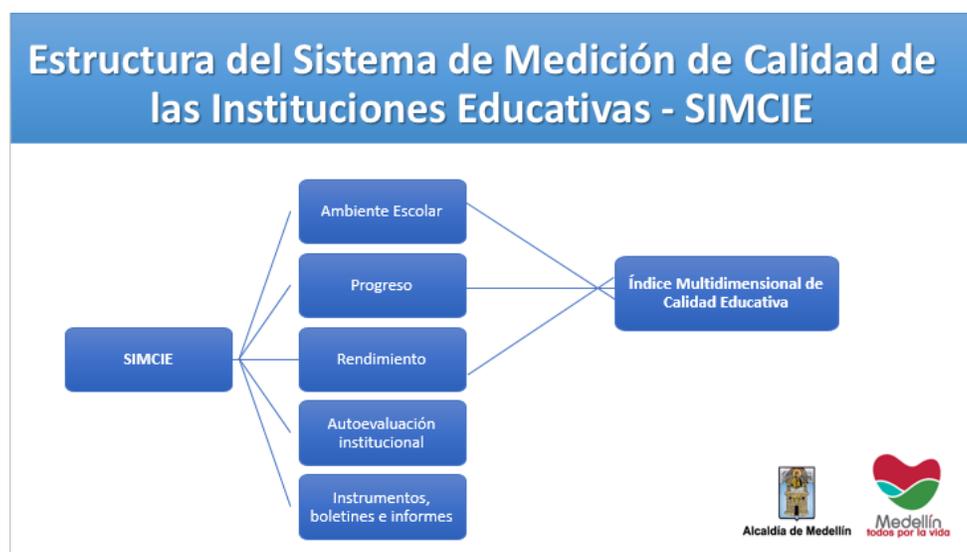
Anderson García

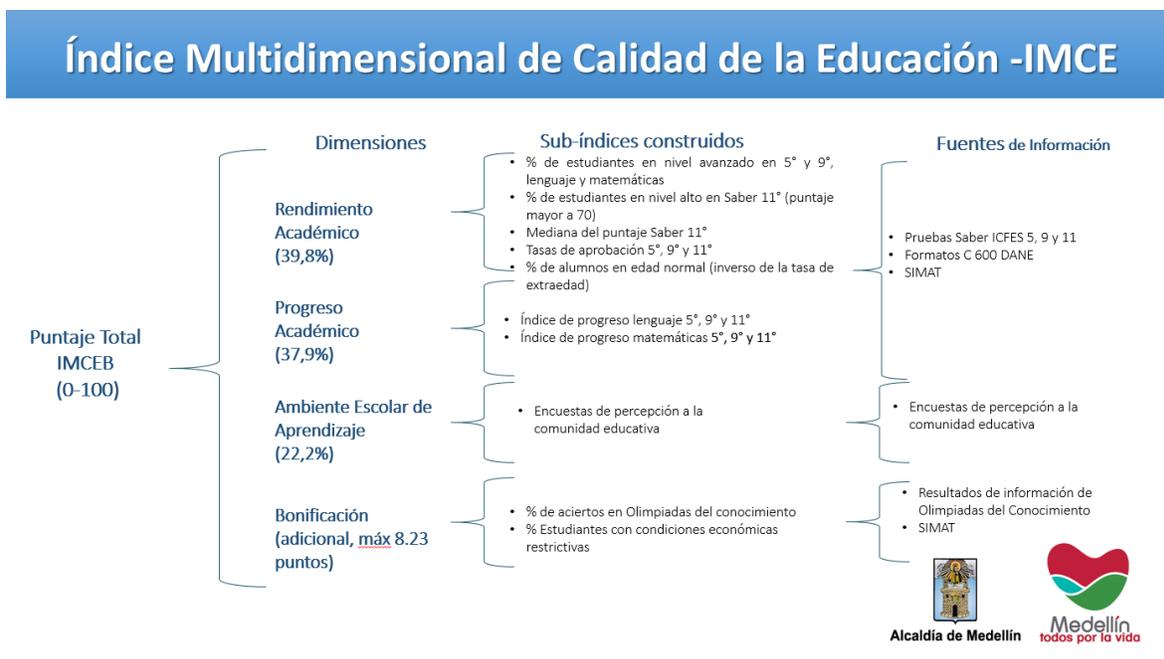
Esta ponencia explica el Sistema de Medición de la Calidad de las Instituciones Educativas (SIMCIE), una herramienta en línea que permite el seguimiento y evaluación del sistema educativo de Medellín. SIMCIE se construye bajo la premisa que, si bien las pruebas estandarizadas son un instrumento indispensable para conocer la eficacia del sistema educativo, no son el único, y es necesario incluir aspectos como el ambiente escolar de aprendizaje, el cual se ha encontrado que ejerce una influencia importante en la calidad de las escuelas y la autoevaluación institucional, para identificar y revisar la coherencia entre la organización de la institución, sus prácticas, el logro de los aprendizajes de todos los estudiantes, las realidades que la circundan y los propósitos transformadores propios del proceso educativo.

El conferencista señala como objetivos relevantes del SIMCIE servir de apoyo al seguimiento y evaluación que dispone la Secretaría de Educación sobre todo el sistema educativo de la ciudad, y permitir el monitoreo de las metas en términos de calidad educativa enmarcadas en el Plan de Desarrollo municipal vigente, así como de los objetivos permanentes del sector educativo orientados al mejoramiento de la eficacia escolar y el aumento de los aprendizajes de los estudiantes. De forma paralela se construyó la propuesta de Autoevaluación Institucional Integrada, que surge del referente del Progress Report de Nueva York, el cual también fue modelo para la construcción del indicador. En la figura 25, puede verse la estructura del Sistema de Medición de Calidad de las instituciones educativas.

El indicador a través del cual SIMCIE muestra sus resultados se denomina índice multidimensional de calidad de la educación (IMCE), el cual sintetiza en un solo valor los resultados de la institución en las diferentes dimensiones que se consideran relevantes para la calidad educativa, a saber: rendimiento académico, progreso académico y ambiente escolar de aprendizaje. Se construye bajo la premisa que la calidad es multidimensional, es decir, está influenciada por una diversidad de factores que se deben tomar en cuenta para su análisis. Su puntaje va de 0 hasta 100. Cada dimensión tiene un peso específico en el índice y la suma de todas ellas corresponde al total. La figura 26 resume los aspectos más relevantes del IMCE, en la que el ponente resalta que las instituciones se clasifican en cinco niveles (A, B, C, D y E), según su puntaje obtenido y respecto a las instituciones del mismo sector. En el grupo A se ubica el 25% de las instituciones que tienen los mayores puntajes.

**Figura 25.** Sistema de Medición de Calidad de las instituciones educativas.



**Figura 26.** Índice multidimensional de calidad de la educación (IMCE).

García señala que durante las primeras mediciones, por falta de información, se tomaron los últimos datos disponibles, por lo que el indicador puede corresponder a los resultados de años anteriores y no necesariamente al año de medición. En 2013 se homogenizaron las fuentes de información para el cálculo del índice con la información del año correspondiente al índice publicado, e igualmente menciona la imposibilidad de comparar los periodos 2011 y 2012 con 2013-2014. Entre los resultados más importantes, García destaca que el promedio de la ciudad es de 50 puntos aproximadamente y que el ambiente escolar alcanza los niveles más altos. Específicamente, en el año 2014, el promedio del IMCEB para la ciudad se ubicó en 49,26 puntos y se observó una mayor desviación estándar en el rendimiento académico en contraste a ambiente escolar, área que presentó menor variabilidad.

Al evaluar las diferencias entre establecimientos por dimensión, se evidencia que las diferencias entre estratos socioeconómicos son más pronunciadas en el índice rendimiento académico, mientras que en el ambiente escolar y progreso se observan instituciones con resultados destacados a pesar de atender poblaciones de niveles socioeconómicos bajos.

El ponente también describe cómo se analiza la dimensión de ambiente escolar. Para la medición de esta dimensión fue necesario construir tres encuestas, una para cada actor de la comunidad educativa: estudiantes, padres y docentes. Su diseño se basó en dos estudios: el Progress Report de la ciudad de Nueva York y la Evaluación del Programa Escuelas de Calidad

para la Equidad y la Convivencia de la Universidad Eafit, la Secretaría de Educación y Proantioquia. Las encuestas están compuestas por preguntas cualitativas que evalúan la percepción del encuestado con respecto a diferentes componentes. Las categorías de respuesta son: "en completo desacuerdo", "en desacuerdo", "de acuerdo" y "completamente de acuerdo". Estas opciones de respuesta son ordinales y se le asignan un valor entre 0 y 10, siendo este último el mejor puntaje posible. Es decir, se obtiene un puntaje de 10 si la respuesta a la pregunta es "completamente de acuerdo", y así sucesivamente.

Para concluir, García resalta que el IMCE es un insumo de información fundamental para el apoyo que da la Secretaría a los procesos de mejoramiento en las instituciones de la ciudad y facilita la identificación de fortalezas y debilidades relacionadas con la calidad de cada establecimiento educativo, además de servir de insumo para la generación de planes de mejoramiento, así como la reducción de las asimetrías de información entre la ciudadanía y el sector. Finalmente, destaca que los usos de este indicador no se limitan a la Secretaría de Educación, pues las instituciones educativas reconocen su importancia y lo han integrado a sus procesos de autoevaluación y mejora continua.

### *Índice de la calidad educativa de Bogotá*

Luis Alfredo Muñoz

Esta ponencia expone la investigación desarrollada para la construcción del índice de calidad de la educación escolar en Bogotá (ICEB), que surge de la disyuntiva de si es mejor tener un índice único que resuma de manera sintética los diferentes elementos constitutivos de la calidad educativa, o es necesario recurrir a un conjunto de indicadores que permitan reconocer la especificidad de cada una de esas dimensiones de manera individual.

El ponente señala que el objetivo general del indicador es ser una herramienta que permita el diagnóstico, seguimiento y mejoramiento continuo del servicio educativo en Bogotá, especialmente el público, pero al mismo tiempo debe responder a unos objetivos específicos: (i) dotar a la Secretaría de Educación de instrumentos de valoración periódica de la calidad del servicio educativo ofrecido por los colegios, las localidades y el distrito; (ii) proporcionar a las instituciones educativas distritales instrumentos para mejorar su gestión y los resultados educativos; (iii) identificar fortalezas y debilidades en la prestación de un servicio educativo de calidad; y (iv) contar con una línea de base que permitiera realizar comparaciones y establecer compromisos de mejoramiento.

La medición se deriva de la multiplicidad de factores, variables, fuentes de información y herramientas estadísticas que fueron considerados en el desarrollo del proyecto de investigación que dieran cuenta de un concepto de calidad; para el cual es importante el desarrollo de

una estructura multinivel, donde el más pequeño es el establecimiento educativo y la mayor cobertura se tiene a nivel distrital.

El proceso de construcción del índice se desarrolló en cinco fases: la estructuración de un marco conceptual, una propuesta inicial (indicadores y estadísticas), un estudio de factores asociados, indicadores y estadísticas, y finalmente el índice de calidad. Muñoz hace hincapié en el estudio de factores asociados, del cual se resumen las principales consideraciones en la tabla 35, y destaca la construcción de índices de resultado haciendo uso de los datos de las pruebas Saber y de la tasa de terminación de cada nivel educativo, con base en el concepto de la calidad de la educación como la capacidad de los establecimientos educativos de promover a todos y cada uno de sus estudiantes, asegurándoles elevados niveles de competencias.

El ICEB consta de cuatro dimensiones –asequibilidad, accesibilidad, adaptabilidad y aceptabilidad– que mide un subíndice respectivamente, en razón a que la política educativa requiere de un conjunto de indicadores que den cuenta de la calidad y equidad del sistema educativo escolar. El ponente destaca que el efecto de las políticas educativas distritales ha sido más visible en la disponibilidad y acceso a la educación escolar, y menos en la calidad educativa, además de observar que los efectos agregados de la política educativa distrital que privilegia el acceso y la disponibilidad en el mejoramiento de la calidad y equidad tienden a disminuir.

**Tabla 35.** Principales consideraciones en el estudio de factores asociados.

<b>Factores Escolares</b>	<b>Factores Locales</b>
<ul style="list-style-type: none"> <li>❑ <b>Insumos:</b> Elementos físicos y materiales puestos a disposición de los procesos educativos, tales como: (i) Recursos Materiales; (ii) Recursos Humanos; (iii) Infraestructura; (iv) Tecnologías de la Información y las Comunicaciones; (v) Subsidios.</li> <li>❑ <b>Procesos:</b> institucionales, escolares y curriculares que garantizan la obtención de los resultados escolares, incluyen factores tales como: (i) el PEI; (ii) el PMI; (iii) Sistema de Evaluación Institucional; (iv) el Clima Escolar; (v) el Liderazgo del rector; (vi) la participación de los padres; (vi) el Calendario Académico, etc.</li> <li>❑ <b>Resultados</b> incluyen los relacionados con el desempeño de los estudiantes en pruebas estandarizadas, al igual que aquellos indicadores que dan cuenta de la capacidad de los establecimientos educativos de retener y promover a sus estudiantes, como son: (i) SABER [5°, 9°, 11°]; (ii) Comprender [3°, 7°]; (iii) Tasa de Repitencia; (iv) Tasa de Deserción.</li> </ul>	<ul style="list-style-type: none"> <li>❑ <b>Contexto familiar:</b> Donde son relevantes el Nivel Socioeconómico de la familia, la esperanza de vida escolar (EVE), el trabajo infantil, el nivel de educación preescolar, genero, y el acceso al Internet.</li> <li>❑ <b>Contexto socio-territorial:</b> factores intervinientes tales como: la distancia a la escuela, el amoblamiento urbano, la calidad de vida, los apoyos institucionales, etc.</li> <li>❑ <b>Contexto escolar:</b> Donde intervienen factores como: el vecindario, el tamaño del colegio, los niveles, las jornadas y los riesgos ambientales.</li> </ul>

### *Sistema de calidad e indicadores de seguimiento para Chile*

Carlos Henríquez Calderón, secretario ejecutivo de la Agencia de Calidad de la Educación de Chile

La conferencia se centra en el Sistema de Aseguramiento de la Calidad e Indicadores de Seguimiento en Chile, el cual surge con el propósito de asegurar la calidad del sistema educativo en su conjunto, estableciendo el desempeño de las escuelas y orientando a los sostenedores, docentes y directivos para el logro de dicho objetivo. El ponente explica las funciones y énfasis de la Agencia de Calidad de la Educación, que es la institución responsable de evaluar y monitorear los logros del aprendizaje y de otros indicadores de calidad, así como de orientar a los establecimientos para avanzar en una educación de calidad para todos y generar información que contribuya a disminuir las brechas de aprendizaje y generar una educación más inclusiva.

Henríquez primeramente contextualiza al público con la situación actual de la educación chilena, remarcando la existencia de una demanda por calidad con equidad que trasciende la discusión especializada, ya que se observa una distribución desigual de oportunidades, como cita Andreas Schleicher, director de educación de la OECD: "En Chile, los estudiantes son segregados desde edades tempranas, lo que refleja la idea de que sólo algunos niños pueden alcanzar estándares de clase mundial, sobre todo los de mejores antecedentes familiares". Luego, existen desafíos importantes en materia de equidad educativa que reflejan que el componente socioeconómico continúa siendo el factor externo a los establecimientos que más explica los resultados del aprendizaje, pero también es relevante que el Estado garantice los pisos mínimos de aprendizaje para todos los estudiantes: 63,1% de los estudiantes no alcanza el nivel adecuado (suma entre nivel elemental e insuficiente). De una forma más explícita, 140.587 estudiantes de un total de 222.800.

El conferencista señala que la calidad educativa es un fenómeno multidimensional y la define como la capacidad del sistema educacional para desarrollar las potencialidades y aprendizajes de todos los estudiantes, como los conocimientos, actitudes y habilidades requeridos en las distintas dimensiones de su formación escolar (que van más allá de lo medido por SIMCE y los OIC), de manera tal que los estudiantes tengan las herramientas que les permitan desarrollarse personalmente y participar, con base en sus preferencias, en los distintos ámbitos de la vida social, política y económica, entre otros. En este contexto, la reforma educacional provee las condiciones para mejorar la calidad de la educación, colocando a los docentes como techo de la calidad de la educación de un país y transformando la educación pública en el referente de la calidad educativa en Chile.

En relación a los indicadores de seguimiento, Henríquez hace referencia a un sistema de evaluación nacional y a la participación de Chile en pruebas internacionales. En cuanto a las pruebas nacionales, el sistema aplica evaluaciones anuales en diferentes asignaturas a todos los

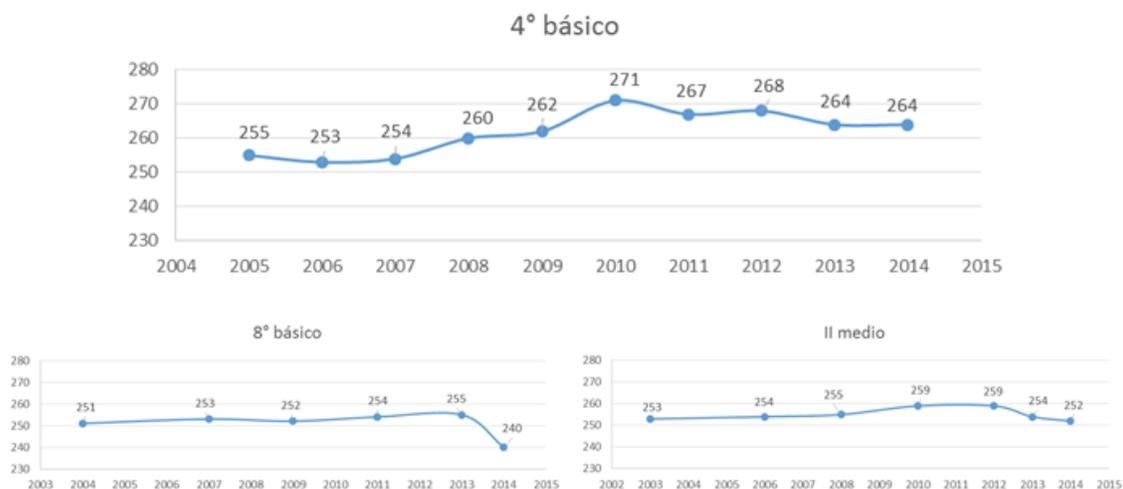
estudiantes del país que cursan los grados evaluados. Hasta 2015, el plan de evaluaciones examinó en seis de los doce grados de la trayectoria escolar y en seis asignaturas; sin embargo, entre 2010 y 2015 aumentó significativamente la cantidad de evaluaciones, lo cual ha generado una importante presión en el sistema escolar.

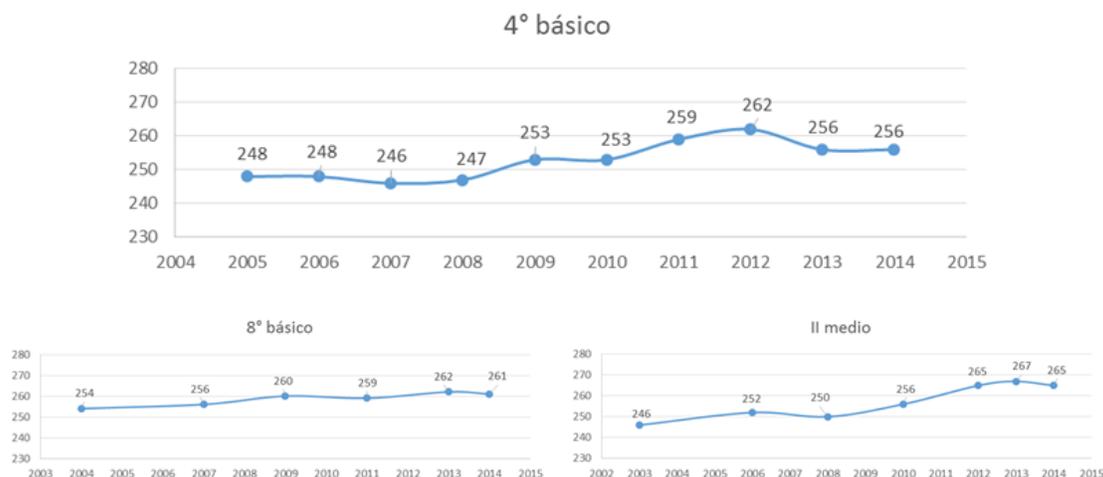
Al examinar cómo han evolucionado los resultados, Henríquez señala que en lectura (figura 27) se ha observado en 4° básico un aumento de los resultados y estabilidad en los últimos años, mientras que en 8° básico se evidencia estabilidad en los resultados hasta el año 2013 y hay una caída de los resultados el año 2014. Esta baja es atípica y presenta desafíos para el país en materia de fomentar los hábitos de lectura de los estudiantes y otras medidas que ayuden a mejorar estos resultados. En el medio observamos estabilidad en el período presentado.

Para matemáticas (figura 28), en general, se observa un crecimiento en los períodos evaluados pero llegado a una meseta que muestra estabilidad. En este sentido se necesitan esfuerzos renovados para avanzar a un mayor logro de aprendizajes en esta asignatura.

En el contexto internacional, Henríquez señala las pruebas (tabla 36) en las que Chile participa y esboza algunos resultados, como: respecto a la media internacional, Chile se ubica debajo de ella en las diferentes áreas y grados evaluados (PISA, TIMSS, ICILS, ICCS); respecto a la media regional, en base al estudio TERCE, Chile se encuentra consistentemente sobre el promedio regional en 3° y 6° grado en todas las áreas evaluadas; y respecto a la comparación regional según al último estudio PISA (2012), Chile se ubica sobre el promedio latinoamericano en todas las áreas evaluadas.

**Figura 27.** Promedios de evaluaciones de lectura.



**Figura 28.** Promedios de evaluaciones de matemáticas.**Tabla 36.** Pruebas internacionales de las que Chile es participante.

Sigla	Nombre del estudio	Qué evalúa	A quiénes evalúa
ECES	Estudio Internacional de Educación en Primera Infancia	Fase de elaboración de perfiles institucionales y políticos de los países	Último año de educación parvularia, antes de comenzar enseñanza obligatoria
ICCS	Estudio Internacional de Educación Cívica y Formación Ciudadana	Educación cívica y formación ciudadana	8° grado
ICILS	Estudio Internacional de Alfabetización Computacional y Manejo de Información	Alfabetización computacional y manejo de información en un contexto digital	8° grado
PIRLS	Estudio Internacional del Progreso en Competencia Lectora	Competencias lectoras y actitudes hacia la lectura	4° grado
TIMSS	Estudio Internacional de Tendencias en Matemática y Ciencias	Matemática y ciencias	4° y 8° grado
PISA	Programa para la Evaluación Internacional de Estudiantes	Lectura, matemática, ciencias (más dominios adicionales por ciclos)	15 años
TERCE	Estudio Regional Comparativo y Explicativo	Lectura, matemática, ciencias y escritura	3° y 6° grado

En relación al desempeño de las escuelas y la evaluación de los procesos de gestión escolar, se establece una clasificación de las instituciones por medio de la construcción de un índice que se basa en los resultados SIMCE y en otros indicadores de la calidad, el cual es ajustado después

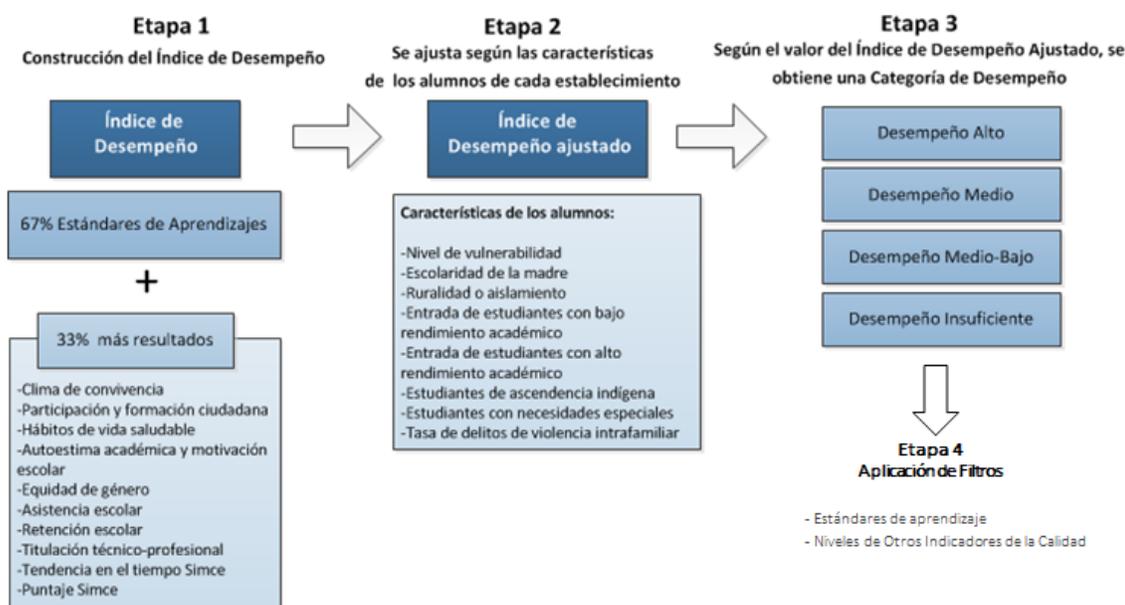
por las variables de características de los alumnos (la estimación de los coeficientes para la corrección se obtiene por medio de una regresión con datos a nivel individual). El índice final corregido corresponde a la parte no explicada por las características de los alumnos, según índice construido en base a resultados SIMCE y otros indicadores de la calidad. Posteriormente se clasifican los establecimientos de acuerdo a la posición de la escuela con respecto a las otras en función del índice final. Aquí es importante señalar que los niveles de corte entre categoría se determinan en la primera clasificación y quedan fijos. Finalmente, se aplican dos filtros de categorías basados en estándares de los indicadores: estándares de aprendizaje y niveles de otros indicadores de la calidad (figura 29).

Este proceso de clasificación permite a los establecimientos conocer su nivel de desempeño y familiarizarse con los dispositivos de apoyo que involucra la marcha blanca del Sistema de Aseguramiento de la Calidad (SAC). Este sistema de clasificación permite establecer las visitas y orientación pertinente, como se encuentra estipulado en las leyes chilenas (tabla 37).

Las visitas constituyen una de las herramientas que tiene la Agencia para evaluar y orientar a los establecimientos educacionales y sus sostenedores, con el objetivo de promover la mejora continua de la calidad de la educación entregada por los establecimientos y fortalecer sus capacidades de autoevaluación en el quehacer institucional. Finalmente concluye nombrando alguno de los desafíos del sistema de educación chileno, como la transformación de los resultados en orientaciones útiles para el mejoramiento de la calidad, el fortalecimiento de las capacidades de autoevaluación de las escuelas, la integración de las distintas realidades escolares en la evaluación y el complementar el diagnóstico interno y externo para retroalimentar los planes de mejoramiento educativo.

**Tabla 37.** Sistema de visitas por nivel de desempeño.

Categoría de desempeño	Frecuencia de visitas
Alto	A solicitar y visita de aprendizaje (máximo el 5% de las visitas anuales)
Medio	Por determinar
Medio bajo	Al menos cada 4 años
Insuficiente	Al menos cada 2 años

**Figura 29.** Construcción del índice final corregido.

### *Sistema de calidad e indicadores de seguimiento para Medellín*

Claudia C. Rave, Gobernación de Antioquia

La Gobernación de Antioquia, en su Plan de Desarrollo “Antioquia, la más Educada”, ha venido construyendo una plataforma integral de gestión de la educación que garantice cobertura, equidad y transparencia del sistema de formación, en un marco de trabajo que sitúe la educación como concepto y propósito transformador en todos los niveles de la sociedad.

Desde el punto de vista del sistema de información, se ha consolidado la Ruta de la Calidad Educativa (figura 30) como un proceso integral que convoca a todos los actores de la sociedad a pensar en la educación como motor de la transformación cultural y que los compromete con la revisión, análisis y construcción de las mejores prácticas para la sostenibilidad de una cultura educativa. La Ruta de la Calidad consiste en la visibilidad pública de todas las apuestas en materia de gestión de la educación, empezando por la firma de pactos y acuerdos públicos en los que se presentan ante la comunidad el gobernador, los alcaldes, los rectores y los maestros, en un evento que sitúa en la primera plana a la educación como estrategia y materia de seguimiento continua. Esta práctica ha tenido profundas repercusiones en Antioquia, en tanto que la ciudadanía puede hacer seguimiento directo a través de los gobernantes que eligió y que reconoce las apuestas educativas de sus hijos y comunidades en un formato que adquiere relevancia política y cultural.

Figura 30. Ruta de la Calidad Educativa.



Durante la exposición se revisó brevemente las actuaciones en todos los niveles de la gestión educativa en Antioquia, enfatizando cómo estas se proyectan para que su comunicación pública tenga eco en la discusión diaria, pública y abierta, de tal manera que construya nuevos imaginarios sociales, políticos y culturales alrededor de la educación. Desde los pactos por la calidad hasta las Olimpiadas del Conocimiento, el propósito de la Ruta de la Calidad ha sido el de movilizar a la sociedad.

Igualmente, la Ruta de la Calidad ha ofrecido herramientas para materializar la posibilidad de seguimiento, evaluación, verificación y debate de todos los proyectos educativos que han propendido por subir los niveles de calidad. Se presenta entonces una discusión sobre cómo resulta el ISCE para nueve colegios de innovación en Antioquia (el programa Colegios de Innovación representa uno de los proyectos de transformación escolar más importantes) y desde allí se realizan análisis, observaciones y recomendaciones a la estructuración del indicador. Ello desde el reconocimiento de que el ISCE es un mecanismo de evaluación importante que requiere una dinámica de seguimiento en el mediano y largo plazo, que le permita capturar progresivamente el valor agregado que los múltiples esfuerzos por mejorar la educación en el país genera para la sociedad.

### *Las voces de los sujetos. Sistema de monitoreo al Plan Sectorial de Educación*

Jorge A. Palacio, Subdirección Académica IDEP

El Plan Bogotá Humana estableció como una de sus metas un sistema de evaluación de la calidad de la educación unificado y de monitoreo al Plan Sectorial de Educación, para lo cual, la Secretaría de Educación (SED) y el Instituto para la Investigación Educativa y el Desarrollo Pedagógico (IDEP) acordaron en 2013 abordar la tarea por partes iguales, asumiendo la SED el desarrollo del Sistema de Evaluación y el IDEP el Sistema de Monitoreo al Plan Sectorial. El IDEP estableció que dicho sistema se diseñaría desde la perspectiva con la que el IDEP venía avanzando en una de sus líneas de investigación, el componente de educación y política pública.

Esta línea de trabajo había considerado el abordaje de los estudios en la comprensión de que los derechos son de las personas y que una mirada desde los sujetos, como titulares de los mismos, aportaría nuevos elementos para la comprensión del derecho a la educación. Esta mirada es diferente a la planteada por Katarina Tomasevski desde su enfoque de las cuatro "A" en el horizonte del cumplimiento de las obligaciones del Estado, que ha sido una vertiente fundamental en los últimos años y ha orientado el trabajo en varias entidades, incluido el IDEP. Un segundo aspecto tuvo que ver con que la mayoría de los estudios se orientaban a la evaluación de la política pública y en este caso lo que interesaba era la valoración por parte de los sujetos. Estos dos conceptos son muy próximos y en muchos contextos parecen indiferenciables. Pero el interés estaba puesto en referentes que dieran cuenta de lo que se percibe en contextos específicos sobre el devenir de las políticas, más que en medidas objetivas referidas a estándares generales de comparación.

Debido a que se encontraban perfeccionando el modelo para el ejercicio de monitoreo, su aplicabilidad y replicabilidad, un diseño emergente y adaptable a las particularidades de las apuestas y/o los proyectos a abordar con respecto al actual PSE era pertinente, pero igualmente importante era su adaptabilidad para el monitoreo de planes sectoriales a futuro. Bajo dicho enfoque se plantea que el diseño "emerge" sobre la marcha, cambiando según el curso de su aplicación y a medida que los investigadores toman decisiones, a partir de lo que han descubierto las vivencias de los actores educativos.

En clara conexión con esta mirada aparece la complementariedad de los enfoques cualitativo y cuantitativo en el ejercicio de monitoreo al PSE (especialmente si se orienta a comprender las vivencias de los sujetos), en la medida en que estos procesos implican tanto aspectos observables como elementos medibles, enunciados como verificables, para en últimas ser interpretados.

En esta ponencia se presentó el rumbo seguido en la construcción del modelo, que dirige el foco hacia la combinación de ejercicios de indagación cualitativa e indagación cuantitativa, y mira con particular atención el ejercicio de pilotaje llevado a cabo en 2015, con especial atención

a la Jornada 40 x 40, de particular interés para la política educativa actual en el contexto distrital y nacional. La figura 31 describe el camino seguido para el monitoreo al Plan Sectorial de Educación de Bogotá.

**Figura 31.** Monitoreo al Plan Sectorial de Educación de Bogotá.



## Conferencias Call for Papers

*Brechas de género en el rendimiento escolar a lo largo de la distribución de puntajes.*

*Evidencia de las pruebas Saber 11*

Luz Karime Abadía y Gloria Bernal Nisperuza, convocatorias Icfes

Acorde con la literatura internacional, la desigualdad en pruebas que determinan el ingreso a la educación superior está relacionada con el menor acceso a oportunidades laborales para las mujeres relativo a hombres con similares habilidades, y con el desarrollo de una sociedad más desigual. El presente trabajo mide la brecha de género en las pruebas Saber 11 de 2014 a lo largo de la distribución de los puntajes. Se encuentra que, para los puntajes globales, matemáticas y lenguaje, la brecha observada es monótonicamente creciente, y que la diferencia se acenúa en contra de las niñas en la cola alta de la distribución. Este comportamiento se mantiene

incluso al controlar por características individuales, familiares y escolares de los estudiantes y al usar diferentes técnicas de estimación. La brecha encontrada, junto con el hecho de que haya menos niñas en los cuantiles altos de la distribución pese a que hay más niñas en el sistema educativo, plantea grandes inquietudes sobre las oportunidades de acceso a la educación superior que actualmente tienen las mujeres frente a los hombres en el país.

### *¿Vale la pena prepararse? Efectos de la participación en actividades preparatorias sobre el desempeño en el examen Saber 11*

Silvia Gómez Soler y Paula Herrera, convocatorias Icfes

La pregunta de investigación de este estudio se centra en determinar si la preparación previa a la presentación de las pruebas Saber 11 genera un impacto positivo sobre el resultado final de las mismas. Adicionalmente, las autoras comparan este impacto en diferentes submuestras y en este caso se concentran en determinar diferencias en el género, en la jornada escolar y en si el colegio es oficial o privado.

El estudio cuenta con entrevistas a rectores y estudiantes de colegio, lo que les permite a las autoras encontrar información anecdótica que complementan los resultados obtenidos. Se estimaron los resultados a través de mínimos cuadros ordinarios como primera aproximación y se controla por variables de interés. Pero dado que las personas que deciden prepararse son diferentes a las que no lo deciden, es necesario que exista una metodología de comparación que en lo único que difieran las unidades de medida sea en la decisión de la preparación, esta metodología se conoce como Propensity Score Matching (PSM).

Dentro de los resultados sobresale que el 73% de los estudiantes deciden prepararse para la prueba y un 71% lo hace en el mismo colegio. La preparación supone un puntaje más alto global de 4 puntos, un efecto bajo.

Los resultados diferenciados por subgrupos sugieren que la preparación tiene un efecto positivo en las mujeres, aunque el valor es muy bajo (0,7 puntos); por otro lado, ser de un colegio público o pertenecer a estratos socioeconómicos bajos genera un efecto negativo sobre el puntaje.

### *Predicción automática de la dificultad de preguntas abiertas de respuesta corta con propósitos educativos*

George Dueñas

El objetivo de este estudio se centra en encontrar un método para la predicción automática de la dificultad de preguntas abiertas. Para contextualizar, en la construcción de preguntas de

evaluación, el evaluador se enfrenta al problema de elegir las preguntas para diferenciar a los estudiantes de acuerdo con su nivel de habilidad o conocimiento.

Se ha visto la necesidad de introducir las preguntas abiertas en las pruebas debido a que estas evalúan capacidades cognitivas diferentes a las evaluadas mediante las preguntas cerradas. La selección de estas preguntas para una prueba implica problemas tales como altos costos y riesgos de confidencialidad. La selección de preguntas apropiadas presupone conocer el nivel de dificultad de cada una. El método empleado se basó en la extracción de 46 factores de los 196 ítems de SciEntsBank. Los factores fueron obtenidos de los textos de la pregunta, la respuesta y la demanda cognitiva. Se obtuvieron 2.070 características de los 46 factores combinados para cada ítem. El método supervisado KBest fue utilizado para seleccionar las mejores características.

La dificultad de cada ítem fue predicha a partir de las mejores características y dos modelos de regresión. Una de las conclusiones importantes de esta investigación fue que el mejor modelo para predecir la dificultad es el empleo de la máquina de vectores de soporte y la característica de conteo de palabras claves.

### *Valor agregado en lengua extranjera de la educación superior en Colombia*

Ramiro Rodríguez Revilla y Diana Carolina López Cuevas

En este estudio se presenta un modelo para medir el valor agregado para la educación superior en Colombia discriminado por áreas de conocimiento de los programas académicos, utilizado para evaluar la calidad del sistema educativo en este nivel la competencia del idioma inglés, teniendo como marco normativo el Plan Nacional Bilingüe 2004-2019 y el Marco Común Europeo de Referencia.

Para llevar a cabo esta medición se utiliza una estimación de modelos lineales jerárquicos con variables instrumentales a partir de un panel de microdatos suministrados por el Icfes y el DANE. El resultado del modelo muestra que los programas de economía, contaduría y administración de empresas son los que más aportan a la formación en inglés en Colombia. El valor agregado de los programas de ciencias de la educación representa el 40% del aporte de los que se encuentran en los primeros lugares de la lista, evidenciando que es necesario fortalecer la estructura curricular de los futuros profesores y licenciados en Colombia. A partir de estos resultados se elaboran algunas recomendaciones de política para mejorar la calidad de la educación en lengua extranjera.

## *Factores asociados al desempeño de los estudiantes de psicología en las pruebas*

### *Saber Pro*

Guillermo Augusto García Méndez

El estudio se centra en establecer la asociación que existe entre algunas características de los programas de psicología en Colombia y algunas variables de caracterización de sus estudiantes, con la ejecución de los mismos en los módulos genéricos de las pruebas Saber Pro en los periodos 2011 a 2013.

La presente investigación contó con tres fases de análisis. Durante la primera etapa del proyecto se realizaron análisis descriptivos y comparativos de las variables de caracterización tanto de los estudiantes como de las IES a que pertenecen, de manera general y discriminando por cohorte de aplicación de la prueba y por zonas geográficas (región y departamento). En segunda instancia se realizó un análisis factorial de correspondencias múltiples, el cual permitió establecer coordenadas factoriales para cada categoría de cada variable, haciendo visibles algunas relaciones entre dichas categorías y los niveles de desempeño de los estudiantes en los módulos genéricos de la prueba. Por último, teniendo en cuenta las estructuras anidadas en las que se enmarca el modelo de educación (estudiantes dentro de instituciones), se realizó un análisis de tipo jerárquico multinivel, el cual, dadas sus características, posibilitó cuantificar (con mayor precisión que en los modelos de regresión tradicionales) el nivel de asociación de una variable o factor a un mejor o peor desarrollo de las competencias de los estudiantes de psicología, dado por el desempeño en los módulos específicos de las pruebas Saber Pro.

Uno de los resultados más relevantes es la caracterización de los estudiantes de psicología que presentaron la prueba, encontrando que la mayoría son mujeres, de estratos 2, 3 y 4, pertenecientes en gran proporción a programas presenciales de la región Bogotá-Cundinamarca. Adicionalmente se encontró que los desempeños altos en la mayoría de componentes se asocian con altos niveles educativos, ocupacionales y salariales del núcleo familiar del estudiante, con un rango de edad entre 19 y 23 años al momento de presentar la prueba, y con características institucionales relacionadas con el estado de acreditación del programa y el costo de la matrícula.

## *Etnoeducación y desempeño escolar en la región pacífica colombiana*

Fabio Sánchez Torres y Glenda Palacio, convocatorias Icfes

Este artículo explora la influencia del modelo etnoeducativo sobre el logro escolar de los estudiantes del Pacífico usando análisis cuantitativo y con alguna evidencia cualitativa que permita contextualizar los resultados. Dentro de la aproximación cuantitativa, y con el objetivo de encontrar efectos causales, se usa un modelo de efectos fijos en el cual se explota la distribución aleatoria de docentes dentro de las IE en cada área de enseñanza.

Para el segundo enfoque se hicieron entrevistas a rectores y coordinadores de instituciones etnoeducadoras para conocer las prácticas educativas que desarrollan al interior del aula escolar. Los resultados econométricos indican que la proporción de docentes etnoeducadores en un área de conocimiento está asociado con mayor logro escolar. Esto lo corrobora la evidencia cualitativa, al mostrar que en las instituciones educativas se incentiva el aprendizaje involucrando a la comunidad educativa (padres de familia, docentes y estudiantes) lo cual motiva los procesos de enseñanza y aprendizaje.

### *Desigualdad de oportunidades en la educación media y superior en Colombia*

Jaime Andrés Sarmiento Espinel y Adriana Carolina Silva Arias, convocatorias Icfes

Además de ser un bien valioso en sí, la educación es un medio de igualar oportunidades entre ciudadanos para que puedan mejorar su bienestar y su nivel de ingresos. No obstante, es limitada la literatura que evalúa el grado de desigualdad educativa presente en Colombia. En este sentido, esta investigación tiene como propósito analizar el grado en el que algunas características de los individuos, que no pueden ser modificadas por sus propias acciones, afectan sus oportunidades de obtener mejores resultados académicos en la educación media y superior en Colombia. Para ello, se analiza la distribución del logro educacional de las personas que al terminar su ciclo de educación media presentaron la prueba Saber 11 entre los años 2000 y 2007, y que en el segundo semestre de 2011 estaban terminando su ciclo de educación superior y presentaron la prueba Saber Pro. De esta forma, se evidencia que a medida que transcurre menos tiempo entre el momento en el que las personas están finalizando la secundaria y la universidad, se vuelven más relevantes las circunstancias para explicar las diferencias en su rendimiento académico.

### *Efectos de largo plazo del Programa Mamá Canguro sobre resultados educativos y de fuerza laboral*

Darío Maldonado

Este estudio evalúa los efectos de largo plazo del Programa Mamá Canguro sobre varios resultados a nivel educativo y laboral. La asignación de los niños a cualquier incubadora o cuidado de la mamá canguro fue aleatoria en razón a que se tenía la ventaja de un experimento aleatorio. Se evidencia que los niños pertenecientes al programa estuvieron más tiempo matriculados en preescolar y tenían un periodo de ausencia temporal menor en el colegio. Un resultado adicional hace referencia a los individuos que ya se encuentran en el mercado laboral, cuyos salarios son más altos en relación a sus contrapartes. Se encuentra, además, que en promedio los niños del cuidado de mamá canguro obtienen peores resultados en los exámenes de lenguaje y matemáticas. Este resultado adverso se concentra en los niños más prematuros. Estos resultados son compatibles con los hallazgos iniciales usando este experimento aleatorio,

que muestra que el programa de cuidado de mamás canguros protege a los niños más vulnerables.

### *Educación media y mano de obra industrial en Colombia. Inequidad y pobreza*

Jhon James Mora Rodríguez

La literatura ha mostrado la estrecha relación existente entre el desarrollo industrial y la acumulación de capital humano, y ha sido precisamente en aquellas regiones del país en las cuales la población exhibe bajos niveles de capital humano las que han tenido grandes dificultades para adaptarse a los procesos de industrialización que se venían desarrollando durante el siglo XX. Este retraso en materia educativa, teniendo claro que la educación promueve una mayor productividad, no permitió un desarrollo homogéneo de la industria moderna del país, pues en las regiones con un mayor nivel de capital humano se desarrollaron procesos de industrialización con mayor innovación.

Precisamente por el impacto que la educación tiene sobre los procesos de industrialización, este artículo considera las ciudades en las cuales la industria nacional concentra la mayor proporción de mano de obra industrial (Bogotá, Cali y Medellín) con el propósito de estudiar cómo sus características sociodemográficas y socioeconómicas inciden en los resultados que en las pruebas Saber 11.

El principal resultado establece que el sistema de educación media en la zona industrial de Colombia promueve la reproducción de una estructura social que es inequitativa en el sentido de que los estudiantes de bajos recursos, al estar concentrados principalmente en las medias jornadas, mañana y tarde, y en una alta proporción con padres que tienen bajos niveles educativos, producen bajos resultados en la prueba Saber 11. Por consiguiente, son estudiantes que tendrán pocas oportunidades para acceder al sistema de educación superior, sea por los bajos ingresos percibidos por sus padres o sea por sus bajas competencias escolares, de tal manera que con una alta probabilidad reproducirán no solo la baja productividad de sus padres, sino también sus bajos ingresos al no acceder a un proceso de formación escolar a nivel superior, quedando atrapados en una trampa de baja productividad y pobreza.

### *¿Qué pasa en el salón de clases con el examen Saber 11? Una etnografía escolar sobre colegios exitosos*

Ana María Zamora Moreno, convocatorias Icfes

Esta ponencia busca responder a la manera como los estudiantes y profesores interpretan y se apropian al examen Saber 11 en la práctica pedagógica. Utilizando los resultados de las pruebas Saber 11 y la clasificación de planteles educativos realizada por el Icfes para el año 2013, se escogieron tres colegios. En estos se realizó un registro etnográfico desde enero hasta agosto de

2014, consistente, entre otras cosas, en siete observaciones de clases de materias presentes en el examen Saber 11 y cuatro observaciones de clases de preparación para el examen. Los audios y observaciones se transcribieron y analizaron con el programa Atlas Ti. Los resultados del análisis de las clases se centran en los momentos en que el examen es nombrado, cuando los estudiantes son entrenados a través de ejercicios para el examen, cuando se les entregan simulacros de la prueba o en discusiones que se generan en las clases pre Icfes.

Las categorías buscan mostrar cómo los estudiantes y el profesor integran el examen a su cotidianidad escolar. Se discuten tres categorías, que son efecto de las prácticas pedagógicas que el examen genera en ciertos momentos: la clausura de discusiones, el surgimiento de instrucciones compartidas, y el silenciamiento y legitimación de contenidos.

Para concluir, el examen Saber 11 en los tres colegios observados modela prácticas pedagógicas temporalmente enmarcadas, unas fuera del aula, como son los pre Icfes, y otras en el espacio escolar. En estas prácticas hay una coexistencia entre el contenido, los debates, las argumentaciones y la vinculación del tema tratado con el entrenamiento para el examen. Los estudiantes terminan viendo solo las pistas que les permite subordinarse a asumir criterios externos de evaluación, criterios que no tienen que ver con el contenido pero que les permite entrar a la dinámica socialmente construida frente al examen de aprender para responder. Aunque esta investigación no se puede generalizar por su carácter etnográfico, es importante resaltar la necesidad de este tipo de estudios en diferentes contextos.

### *Desigualdad de oportunidades en Colombia. Impacto del origen social sobre el desempeño académico y los ingresos de graduados universitarios*

Andrea Cuenca, convocatorias Icfes

A partir del cruce de las bases Saber 11, Saber Pro y OLE se realizó un análisis estadístico del impacto del origen social (OS) sobre los logros educativos y ocupacionales de graduados de educación superior en Colombia (n=16,899). Se distinguieron los niveles socioeconómico y educativo de las familias como dos factores mediante los cuales el OS determina: (i) el desempeño académico, (ii) el tipo de institución educativa y (iii) los ingresos en el mercado laboral formal. Para modelar los logros educativos y ocupacionales se realiza un *path analysis* con técnicas de regresión lineal múltiple. Los resultados sugieren que el OS tiene un impacto significativo sobre los logros en la educación media, durante la universidad y la posterior inserción al trabajo. En particular indican que el OS opera fuertemente, de manera directa a través del nivel socioeconómico y de la escolaridad de los padres, sobre el desempeño en la educación media. El OS también opera indirectamente, a través de la segmentación de trayectorias educativas, sobre el desempeño en la universidad y los ingresos.

### *Convergencia espacial del rendimiento educativo en Colombia (2000-2012)*

Osmar Leandro Loaiza Quintero y Guillermo David Hincapié Vélez

El propósito de esta investigación consiste en determinar las brechas regionales a nivel municipal en calidad educativa en Colombia, siendo medida la calidad a través de las pruebas Saber 11 en el periodo 2007-2012. También se midió indirectamente el efecto del vecindario (o la asociación espacial) sobre la probabilidad que tiene un municipio de escalar y descender en la distribución del rendimiento educativo municipal. Se utilizaron los resultados a nivel municipal de las pruebas Saber 11 como una medida de la calidad de la educación. Esta prueba estandarizada comprende seis áreas básicas: biología, química, física, filosofía, matemática y lenguaje.

Las herramientas utilizadas en el desarrollo de la investigación comprende Kernel estocásticos, AEDE univariado y bivariado, matrices de Markov espacialmente condicionadas y modelos de regresión lineal. Los resultados aportados por las Kernel estocásticas, las univariadas y las matrices de Markov o de probabilidad no aportan evidencia que soporte la hipótesis de convergencia intermunicipal en materia de rendimiento educativo. En general se observa que las distribuciones Kernel se desplazan hacia la izquierda o permanecen casi estables, o en su defecto presentan un achatamiento que manifiesta un aumento en la dispersión de los datos. Asimismo, aunque las matrices de Markov presentan un panorama de alta movilidad en las partes medias de la distribución, se observa también que las probabilidades de permanencia en los cuantiles extremos (es decir, de muy bajo o muy alto rendimiento educativo), son relativamente altas, principalmente al considerar el segundo subperiodo de análisis (y excepción hecha de la prueba de filosofía, en donde esto solo es cierto para el primer cuantil). Lo anterior se puede interpretar como que las mayores barreras a la movilidad están en los extremos, no en el centro de la distribución.

### Talleres

Como parte de las actividades del seminario, los días 3 y 4 de noviembre se desarrollaron talleres orientados a familiarizar a los asistentes con metodologías frecuentemente utilizadas en análisis de calidad educativa y con los datos que el Icfes pone a disposición de la comunidad para el desarrollo de actividades de investigación.

Los objetivos de estos talleres fueron: (i) construir y fortalecer las capacidades del país para que estudiantes y grupos adelanten investigaciones sobre la calidad de la educación, (ii) promover el uso de resultados de evaluaciones nacionales e internacionales en investigaciones rigurosas que aporten información confiable para orientar la toma de decisiones sobre mejora de la calidad de la educación, y (iii) propiciar el vínculo entre investigación, práctica educativa y política pública mediante estrategias de divulgación de los resultados obtenidos a través de estos análisis.

En la edición 2015 del seminario se ofrecieron cuatro talleres en relación a las siguientes áreas y de los cuales se dará una breve descripción:

### *Introducción al análisis estadístico de relaciones causales*

Andrés Gutiérrez, director del Departamento de Estadística y Evaluación del Icfes

“¿Qué hubiera sucedido?” es la pregunta que todo investigador debe hacerse si quiere profundizar acerca de las posibles causas de un evento de interés. La realidad de este tipo de investigaciones está supeditada a que sólo es posible observar una realización potencial de la población objetivo, influenciada o no por una intervención de interés. En la investigación de políticas educativas y su posible impacto sobre poblaciones de interés es necesario contar con elementos estadísticos y econométricos confiables que permitan establecer el grado en el cual los factores de interés (comúnmente en forma de intervenciones gubernamentales) afectan positiva o negativamente el comportamiento promedio de las poblaciones de interés. El Icfes, como entidad que produce información relevante de la calidad de la educación en Colombia, cuenta con bases de datos oficiales que investigadores y estudiantes pueden utilizar para establecer relaciones causales entre políticas públicas y el desempeño cognitivo de los estudiantes, el desempeño de los establecimientos educativos y las variables asociadas al logro académico a nivel agregado. El contenido abarcado en el taller incluyó las siguientes temáticas:

- Introducción al modelo causal
- Diseño de experimentos (asignación aleatorizada)
- Estudios observacionales (asignación no aleatoria)
- Métodos de emparejamiento y estimación del puntaje de propensión (*propensity score*)
- Dobles diferencias y datos panel
- Regresión discontinua
- Variables instrumentales
- Métodos de ecuaciones estructurales

### *Introducción al uso de modelos estructurales usando MPLUS*

Lina Saldarriaga, miembro del Comité Asesor del Icfes

El objetivo de este taller fue presentar a los participantes los principales conceptos relacionados con los modelos de ecuaciones estructurales, así como también ayudarlos a familiarizarse con los procedimientos necesarios para crear este tipo de modelos. Se trabajó en el proceso de formulación, identificación, desarrollo y evaluación de modelos usando las bases de datos disponibles en el Icfes. El taller fue teórico-práctico y se desarrollaron ejemplos de algunas de las más

importantes aproximaciones de las ecuaciones estructurales. Todos los ejercicios se desarrollaron usando la versión libre del programa (M-Plus).

### *Introducción a la evaluación de impacto*

Luis Carlos Reyes, profesor asistente de la Universidad Javeriana y miembro del Comité Asesor del Icfes

Este curso cubrió los principios básicos de la evaluación de impacto. Se enfatizó en la importancia de trabajar conjuntamente y desde el inicio de los programas evaluados con las entidades implementadoras, y cómo aprovechar esta dinámica para realizar evaluaciones confiables y de alta calidad. También se abarcaron algunas de las estrategias a seguir cuando el evaluador debe hacer análisis retrospectivos de los datos. Las técnicas cubiertas son aplicables al análisis de políticas educativas, pero pueden ser usadas igualmente en el análisis de programas de salud, microfinanzas y, en general, cualquier tipo de política social.

### *Uso de base datos*

Carolina Lopera, Sandra Roperó y Lorena Wilches, Subdirección Estadística del Icfes

En el desarrollo de este taller se discutió las complejidades metodológicas, implicaciones en el análisis, además de métodos y procedimientos cuantitativos y estadísticos para el uso de las bases de datos proporcionadas por el Icfes.

## IV

### Los participantes y su evaluación sobre el seminario

En el 6° Seminario Internacional de Investigación sobre Calidad de la Educación sobre Índices de Calidad 2015 participaron más de 544 personas en los auditorios del Hotel Marriott de Bogotá. Las conferencias principales y alternas se pudieron seguir vía streaming en todo el país.

En las evaluaciones sobre el seminario, los participantes fueron consultados si el evento les aportaba “nuevos elementos conceptuales y/o metodológicos útiles” para su desempeño académico y profesional, con lo cual coincidió el 89% de los participantes que diligenciaron la evaluación (tabla 38).

Al ser consultados sobre su interés por participar en siguientes versiones del seminario o en eventos similares organizados por el Icfes, el 89% de los participantes manifestó su interés en hacerlo (tabla 39).

De la misma manera, el 65% de los participantes que diligenciaron el formato de evaluación del seminario están de acuerdo con que las temáticas tratadas en este los motivan a hacer investigaciones o seguir investigando sobre temas relacionados con la calidad de la educación (tabla 40).

A su vez, el 84% de los participantes que diligenciaron la evaluación estuvieron de acuerdo en que “las temáticas de las conferencias presentadas son pertinentes respecto a la actualidad de la política educativa” (tabla 41).

Adicionalmente, el 89% considera que “a partir de los resultados de las investigaciones presentadas pueden tomarse acciones para impactar positivamente la calidad de la educación” (tabla 42).

**Tabla 38.** ¿Me aportó nuevos elementos útiles para mi desempeño académico y profesional?

¿Me aportó nuevos elementos útiles para mi desempeño académico y profesional?	2010	2011	2012	2013	2014	2015
Totalmente de acuerdo	60%	76%	67%	64%	66%	67%
Parcialmente de acuerdo	34%	19%	23%	24%	25%	22%
Parcialmente en desacuerdo	2%	2%	3%	3%	3%	4%
Totalmente en desacuerdo	4%	3%	7%	9%	6%	6%

**Tabla 39.** ¿Me gustaría participar en siguientes seminarios?

¿Me gustaría participar en siguientes seminarios?	2010	2011	2012	2013	2014	2015
Totalmente de acuerdo	95%	93%	89%	92%	93%	89%
Parcialmente de acuerdo	2%	5%	3%	3%	2%	3%
Parcialmente en desacuerdo	0%	1%	0%	0%	0%	0%
Totalmente en desacuerdo	3%	1%	8%	5%	5%	8%

**Tabla 40.** ¿Me motiva a investigar sobre calidad de la educación?

¿Me motiva a investigar sobre calidad de la educación?	2010	2011	2012	2013	2014	2015
Totalmente de acuerdo	79%	79%	78%	82%	86%	65%
Parcialmente de acuerdo	17%	18%	13%	12%	9%	19%
Parcialmente en desacuerdo	1%	1%	1%	1%	0%	8%
Totalmente en desacuerdo	3%	2%	8%	5%	5%	8%

**Tabla 41.** ¿Las temáticas son pertinentes a la actualidad de la política educativa?

¿Las temáticas son pertinentes a la actualidad de la política educativa?	2012	2013	2014	2015
Totalmente de acuerdo	72%	62%	60%	65%
Parcialmente de acuerdo	22%	33%	29%	19%
Parcialmente en desacuerdo	2%	2%	6%	8%
Totalmente en desacuerdo	4%	3%	5%	8%

**Tabla 42.** ¿Los resultados puede impactar decisiones política educativa?

¿Los resultados puede impactar decisiones política educativa?	2013	2014	2015
Totalmente de acuerdo	58%	55%	56%
Parcialmente de acuerdo	33%	34%	30%
Parcialmente en desacuerdo	5%	8%	6%
Totalmente en desacuerdo	4%	3%	7%

Al responder la pregunta abierta sobre los aspectos positivos del seminario, las respuestas predominantes subrayan la pertinencia de las discusiones para la política educativa del país, el rigor de los estudios presentados y la continuidad en la realización del seminario.

En cuanto a los temas por mejorar, se subrayan la importancia de la organización física del espacio dada la cantidad de participantes, la necesidad de incluir un enfoque territorial con posibles eventos en regiones del país, la relevancia de contar con asistentes de zonas rurales y la selección de las conferencias nacionales. Los posters son resúmenes de estudios que emplean datos de pruebas estandarizadas de logro utilizados en el seminario por primera vez en 2011, pero solo hasta 2012 fueron evaluados por los participantes. Este año se premiaron los dos mejores poster presentados a través de la votación del público asistente.

Para el 94% de los participantes, los posters son una modalidad de presentación interesante, novedosa y útil (tabla 43). Según el 80% de los participantes, los posters contenían información clara y suficiente. Para el 92% de los participantes, los posters dan una idea del tipo de investigaciones que se pueden hacer con las bases de datos del Icfes.

**Tabla 43.** Percepción sobre póster.

Percepción sobre póster	Novedoso y útil				Información clara y suficiente				Muestra ejemplos			
	2012	2013	2014	2015	2012	2013	2014	2015	2012	2013	2014	2015
Totalmente de acuerdo	64%	60%	70%	63%	49%	57%	65%	53%	72%	69%	69%	69%
Parcialmente de acuerdo	27%	33%	25%	31%	41%	37%	30%	27%	20%	27%	27%	23%
Parcialmente en desacuerdo	5%	5%	3%	3%	6%	6%	3%	15%	20%	2%	2%	5%
Totalmente en desacuerdo	4%	2%	2%	4%	4%	0%	2%	5%	6%	2%	2%	3%

## V

# Conclusiones

En la sexta versión del evento, el Seminario de Investigación sobre Calidad de la Educación sigue posicionándose como un punto de encuentro entre la comunidad académica y los diseñadores de políticas educativas. La importancia que ha adquirido puede verse a través de la asistencia, en persona y vía web; en el número de trabajos de investigación que se reciben anualmente; y en la asistencia a los talleres. El 89% de los participantes considera que el seminario le aportó nuevos elementos útiles para su desempeño académico y profesional y el 65% consideran que las temáticas tratadas los motivan a hacer o continuar investigaciones sobre la calidad de la educación.

Desde su primera edición ha sido un propósito firme de este seminario promover el diálogo entre académicos, comunidad educativa y encargados de la política pública. Esto se ve reflejado en la diversidad del público asistente –investigadores, docentes, estudiantes universitarios, miembros del Ministerio de Educación y las secretarías de Educación– y en las actividades y formatos de presentación. En la presente edición, es la primera vez que se hace hincapié en un tema específico, *los índices de calidad*.

El seminario se desarrolló inicialmente a través del concepto de rendición de cuentas introducido por Susanna Loeb, quien contextualizó al público sobre el debate en torno a los sistemas de rendición de cuentas a través del uso de pruebas estandarizadas a nivel de escuela en el contexto de Estados Unidos, resaltando su utilidad para los hacedores de política, cuyas metas deben ser alineadas a los objetivos propuestos por el sistema educativo e identificar las necesidades que este posee para tomar decisiones de la forma más acertada. De forma subsecuente, Morgan Polikoff presentó desde una perspectiva más técnica la rendición de cuentas en el contexto estadounidense, enfocándose principalmente en el concepto de alineación de la enseñanza. Finalmente Thomas Ahn expuso un estudio que evalúa si las sanciones impuestas por los sistemas de responsabilidad educativa son efectivas. Los resultados de estas investigaciones sugieren que se pueden usar herramientas para comparar los documentos de política de interés (evaluaciones, estándares, materiales curriculares, etc.).

Posteriormente se trataron experiencias tanto nacionales (Medellín y Bogotá) como internacionales (Argentina y Brasil) centradas en índices de calidad. Silvia Montoya describió los

principales retos y oportunidades de los índices compuestos que sirven a los objetivos de desarrollo sostenible, así como la construcción del índice de equidad y calidad de la educación IECEP en Buenos Aires con el objetivo de monitorear mejorar en la escuela, la comuna y el distrito escolar. Andre dos Santos expuso las características del índice de desarrollo de la educación básica (IDEB) para Brasil, el mapa de la calidad de la educación, la evaluación en el Plan Nacional de Educación y los desafíos y perspectivas. Esto fue complementado por Silvana Godoy y Jorge Duarte, quienes presentaron el proceso de desarrollo y la utilidad del índice sintético de calidad educativa en el contexto nacional, y por Anderson García, quien expuso la experiencia de Medellín en la medición de la calidad educativa. Estas conferencias muestran el interés particular de los hacedores de política por crear sistemas de seguimiento y monitoreo para las instituciones educativas, que arrojan resultados visibles al público a través del diseño y construcción de indicadores.

En el segundo día, las ponencias referentes a índices de calidad estuvieron complementadas por las presentaciones del índice de la calidad educativa de Bogotá (ICEB), en la que Luis Alfredo Muñoz explicó la investigación desarrollada para su construcción; luego Carlos Henríquez presentó el sistema de calidad y los indicadores de seguimiento para Chile, contextualizando al público sobre el contexto educativo chileno. Se finalizó con las presentaciones de Claudia Rave y Jorge Palacio. La primera expuso los esfuerzos de la gobernación de Antioquia en la construcción una plataforma integral de gestión de la educación que garantice cobertura, equidad y transparencia del sistema de formación, en un marco de trabajo que sitúe la educación como concepto y propósito transformador en todos los niveles de la sociedad. Palacio presentó un sistema de monitoreo al Plan Sectorial de Educación.

En la sesión de la mañana del segundo día del evento, el énfasis estuvo en las medidas de valor agregado. Jesse Cuhna explicó las posibilidades y limitaciones del uso de datos administrativos para estimar el valor agregado en educación superior en el contexto estadounidense, que a diferencia del sistema educativo colombiano no tiene un test que evalúe el desempeño de los estudiantes al final de su proceso universitario. Seguidamente, Peter Tymss abordó el impacto a largo plazo del acceso de los niños a años sucesivos de educación de alta calidad en las escuelas. Juan Esteban Carranza analizó el crecimiento de la cobertura de la educación superior en Colombia, el cual ha sido espectacular en los últimos quince años, y Ernesto San Martín presentó cómo comunicar indicadores, análisis que se realizó con base en las pruebas Saber 11 y Saber Pro.

## Referencias

- Ahn, T., & Vidgor, J. (2015). The Impact of No Child Left Behind's Accountability Sanctions on School Performance: Regression Discontinuity Evidence from North Carolina. *NBER Working Paper Series*, 3(5), 48. <http://doi.org/10.3386/w20511>.
- Arias Ortiz, E., & Cristia, J. (2014). El BID y la tecnología para mejorar el aprendizaje: ¿Cómo promover programas efectivos?, *Banco Interamericano de Desarrollo*, 75. Retrieved from [http://publications.iadb.org/bitstream/handle/11319/6550/El BID y la tecnolog%C3%ADa para mejorar el aprendizaje%3a %C2%BFC%C3%B3mo promover programas efectivos%3f.pdf?sequence=1](http://publications.iadb.org/bitstream/handle/11319/6550/El%20BID%20y%20la%20tecnolog%C3%ADa%20para%20mejorar%20el%20aprendizaje%3a%20%C2%BFC%C3%B3mo%20promover%20programas%20efectivos%3f.pdf?sequence=1).
- Assuring Quality in Blended Fertilizers*. (n.d.).
- Carey, K., & Manwaring, R. (2011). Growth Models and Accountability: A Recipe for Remaking ESEA. *Education Sector Reports*. .... Retrieved from <http://www.siacabinetreport.com/admin/uploads/articleresources/5252011811502890.pdf>.
- Cunha, J. M., & Miller, T. (2014). Measuring value-added in higher education: Possibilities and limitations in the use of administrative data. *Economics of Education Review*, 42, 64–77. <http://doi.org/10.1016/j.econedurev.2014.06.001>.
- Dee, T., & Jacob, B. (2009). The Impact of No Child Left Behind on Student Achievement. *Journal of Policy Analysis and Management*, 76. <http://doi.org/10.1002/pam>.
- Figlio, D., & Loeb, S. (2011). *School Accountability. Handbook of the Economics of Education* (1st ed., Vol. 3). Elsevier B.V. <http://doi.org/10.1016/B978-0-444-53429-3.00008-9>.
- Grissom, J. A., Kalogrides, D., & Loeb, S. (2015). Using Student Test Scores to Measure Principal Performance. *Educational Evaluation and Policy Analysis*, 37(1), 3–28. <http://doi.org/10.3102/0162373714523831>.
- Guo, J., Marsh, H. W., Morin, A. J. S., Parker, P., & Kaur, G. (2015). Directionality of the Associations of High School Expectancy-Value, Aspirations and Attainment over Eight Years: A Multiwave, Longitudinal study. *American Educational Research Journal*, 52(2), 371–402. <http://doi.org/10.3102/0002831214565786>.

- Hanushek, E. A., Link, S., & Wößmann, L. (2013). Does school autonomy make sense everywhere? Panel estimates from PISA. *Journal of Development Economics*, *104*, 212–232. <http://doi.org/10.1016/j.jdeveco.2012.08.002>.
- Hanushek, E. A., & Rivkin, S. G. (2011). The Distribution of Teacher Quality and Implications for Policy. *Annual Review of Economics*, *4*(1), 131–157. <http://doi.org/10.1146/annurev-economics-080511-111001>.
- Hanushek, E. a., & Woessmann, L. (2012). Do better schools lead to more growth? Cognitive skills, economic outcomes, and causation. *Journal of Economic Growth*, *17*(4), 267–321. <http://doi.org/10.1007/s10887-012-9081-x>.
- Harris, D. N., Ingle, W. K., & Rutledge, S. A. (2014). How Teacher Evaluation Methods Matter for Accountability: A Comparative Analysis of Teacher Effectiveness Ratings by Principals and Teacher Value-Added Measures. *American Educational Research Journal*, *51*(1), 73–112. <http://doi.org/10.3102/0002831213517130>.
- Loeb, S., Soland, J., & Fox, L. (2014). Is a Good Teacher a Good Teacher for All? Comparing Value-Added of Teachers With Their English Learners and Non-English Learners. *Educational Evaluation and Policy Analysis*, *36*(4), 457–475. <http://doi.org/10.3102/0162373714527788>.
- Pincoffs, E. L. (1973). *Educational accountability*. *Studies in Philosophy and Education* (Vol. 8). <http://doi.org/10.1007/BF00372438>.
- Polikoff, M. S. (2012). Instructional Alignment under No Child Left Behind. *American Journal of Education*, *118*(3), 341–368. <http://doi.org/10.1086/664773>.
- Polikoff, M. S. (2015). The Stability of Observational and Student Survey Measures of Teaching Effectiveness. *American Journal of Education*, *121*(2), 183–212. <http://doi.org/10.1086/679390>.
- Polikoff, M. S., McEachin, A. J., Wrabel, S. L., & Duque, M. (2013). The Waive of the Future? School Accountability in the Waiver Era. *Educational Researcher*, *43*(1), 45–54. <http://doi.org/10.3102/0013189X13517137>.
- Polikoff, M. S., & Porter, A. C. (2014). Instructional Alignment as a Measure of Teaching Quality. *Educational Evaluation and Policy Analysis*, *36*(4), 399–416. <http://doi.org/10.3102/0162373714531851>.
- Russell, J. (2013). Epa aape. *Education Policy Analysis Archives*, *21*(12), 1–25. Retrieved from <http://epaa.asu.edu/ojs/article/view/1044>.
- Saavedra, A. R., & Saavedra, J. E. (2011). Do colleges cultivate critical thinking, problem solving, writing and interpersonal skills? *Economics of Education Review*, *30*(6), 1516–1526. <http://doi.org/10.1016/j.econedurev.2011.08.006>.

- Sahlberg, P. (2010). Rethinking accountability in a knowledge society. *Journal of Educational Change*, 11(1), 45–61. <http://doi.org/10.1007/s10833-008-9098-2>.
- Tymms, P., Jones, P., Albone, S., & Henderson, B. (2009). The first seven years at school. *Educational Assessment, Evaluation and Accountability*, 21(1), 67–80. <http://doi.org/10.1007/s11092-008-9066-7>.
- Zambrano, M. F. (2015). Alcances e inconsistencias del Índice Sintético de Calidad Educativa (ISCE), diseñado y aplicado por el Ministerio de Educación de Colombia, 1–17.

Juan Manuel Santos Calderón  
**Presidente de la República**

Gina María Parody d'Echeona  
**Ministra de Educación Nacional**

Víctor Javier Saavedra Mercado  
**Viceministro de Educación Preescolar, Básica  
y Media**

## **Icfes**

Ximena Dueñas Herrera  
**Directora General**

María Sofía Arango Arango  
**Secretaria General**

Hugo Andrés Gutiérrez Rojas  
**Director de Evaluación**

Francia María del Pilar Jiménez Franco  
**Directora de Producción y Operaciones**

Ingrid Picón Carrascal  
**Directora de Tecnología e Información**

Jimmy Alexander Lugo  
**Asesor Oficina Asesora de Comunicaciones y  
Mercadeo**

Luis Eduardo Jaramillo  
**Asesor Dirección General  
Oficina Gestión de Proyectos de Investigación**

Javier Toro  
**Subdirector de Producción de Instrumentos**

Giovany Babativa  
**Asesor Subdirección de Diseño de  
Instrumentos**

Silvana Godoy Mateus  
**Subdirectora de Análisis y Divulgación**

Bogotá, agosto de 2016