

Isáziga-David, Carlos-Hernán; Gabalán-Coello, Jesús; Vásquez-Rizo, Fredy-Eduardo  
La intervención académica en la construcción de una sociedad con calidad: análisis del valor agregado en el  
proceso formativo colombiano  
Hallazgos, vol. 11, núm. 22, 2014, pp. 359-384  
Universidad Santo Tomás  
Bogotá, Colombia

Disponible en: <http://www.redalyc.org/articulo.oa?id=413834075018>



*Hallazgos*,  
ISSN (Versión impresa): 1794-3841  
[revistahallazgos@usantotomas.edu.co](mailto:revistahallazgos@usantotomas.edu.co)  
Universidad Santo Tomás  
Colombia

# La intervención académica en la construcción de una sociedad con calidad: análisis del valor agregado en el proceso formativo colombiano\*

*Carlos-Hernán Isáziga-David\*\**, *Jesús Gabalán-Coello\*\*\**, *Fredy-Eduardo Vásquez-Rizo\*\*\*\**

## RESUMEN

Recibido: 10 de junio de 2014

Evaluated: 28 de junio de 2014

Aceptado: 1 de agosto de 2014

Este artículo presenta una propuesta metodológica basada en la prueba de porcentaje de efectividad máxima, tendiente a estimar el valor agregado existente en la diferencia entre los desempeños relativos de los estudiantes colombianos antes y después de un proceso formativo, tomando como referente las pruebas estatales Saber11 y SaberPRO, y haciendo un análisis comparativo por regiones. Se pone de manifiesto una serie de situaciones que evidencian la gran heterogeneidad, y en muchos casos disparidad, de las regiones y de los departamentos que las conforman, especialmente en el desarrollo de las competencias interpretativas, argumentativas y propositivas de sus estudiantes, con la intención de analizar la posible existencia o no de un nivel académico estudiantil predestinado, asociado a dichas pruebas.

**Palabras clave:** pruebas censales, rendimiento académico, calidad, intervención académica, valor agregado.

\* Esta investigación recibió apoyo del Instituto Colombiano para la Evaluación de la Educación (ICFES). Las opiniones, tesis y argumentos expresados son de propiedad exclusiva de los autores y no representan el punto de vista del instituto. Cómo citar este artículo: Isáziga-David, C.H., Gabalán-Coello, J. y Vásquez-Rizo, F.E. (2014). La intervención académica en la construcción de una sociedad con calidad: análisis del valor agregado en el proceso formativo colombiano. *Hallazgos*, 11(22), 359-384.

\*\* Magíster en Economía de la Universidad Autónoma de Occidente. Correo electrónico: chisaziga@uao.edu.co

\*\*\* Magíster en Ingeniería con énfasis en Ingeniería Industrial de la Universidad Autónoma de Occidente. E-mail: jgabalán@uao.edu.co

\*\*\*\* Magíster en Ciencias de la Información y Administración del Conocimiento de la Universidad Autónoma de Occidente. E-mail: fvasquez@uao.edu.co

## Academic Intervention in the Construction of a Society with Quality: Analysis of the Added Value in the Colombian Educational Process

### ABSTRACT

This article presents a methodological proposal based on the percentage of maximum effectiveness test, aimed at estimating the added value existing in the difference between the Colombian students' relative performances before and after an educational process, taking as a reference the state tests Saber11 and SaberPRO, making a comparative analysis by region. The amount of situations that demonstrate the high heterogeneity, and disparity in many cases, in the regions and departments that make them up, especially in the development of students' interpretative, argumentative and propositive competences are highlighted with the purpose of analyzing the possible existence or inexistence of a predestined academic level, related to these tests.

**Keywords:** Census tests, academic performance, quality, academic intervention, value added.

## A intervenção acadêmica na construção de uma sociedade com qualidade: análise do valor acrescentado no processo de formativo colombiano.

### RESUMO

Este artigo apresenta uma proposta metodológica baseada na prova de porcentagem de efetividade máxima, tendente para estimular o valor acrescentado existente na diferença entre os desempenhos relativos dos estudantes colombianos antes e depois de um processo de formação, pegando como referente às provas do Estado Saber11 e SaberPRO, e fazendo uma análise comparativa por regiões. Põe-se de manifesto uma série de situações que evidenciam a grande heterogeneidade, e em muitos dos casos disparidade, das regiões e dos departamentos que as compõem, especialmente no desenvolvimento de habilidades interpretativas, argumentativas e proposital de seus alunos, com a intenção para analisar a possível existência ou não de um nível acadêmico dos alunos predestinado, associada a tais testes.

**Palavras-chave:** Testes de recenseamento, desempenho acadêmico, qualidade, intervenção acadêmica, valor acrescentado.

## INTRODUCCIÓN

Contar con una sociedad que ostente altos índices de calidad académica pasa necesariamente por una serie de fases en los procesos de formación de los sujetos que la componen. Está ampliamente documentado que los niveles de desarrollo de los países dependen en gran medida de los procesos formativos y de investigación que se llevan a cabo, como parte de una política educativa que pretenda ser sólida. En este sentido, cada uno de los niveles en los cuales se soporta el sistema educativo añade nuevo conocimiento y plantea nuevos y complejos retos, tanto al educando como a las instituciones encargadas de su formación.

La educación superior, desde la perspectiva de la educación permanente, es la etapa en la cual el estudiante experimenta una ruptura con el proceso iniciado en el colegio, dado que en el marco de la autonomía institucional universitaria se desarrollan los mismos programas de formación, que dependiendo del “sello” institucional privilegian unas competencias más que otras. Esto genera implícitamente en la persona el desarrollo de capacidades y apropiación de conocimientos básicos o de profundización para su competente desempeño profesional.

Con base en lo anterior, la pretensión de este trabajo consiste en abordar una parte de la complejidad que involucra la intrínseca formación en el sistema educativo universitario, siendo responsable de la consolidación del propósito de contribuir a la configuración de personas, ciudadanos y profesionales, que en su interacción con el entorno trabajan en beneficio del desarrollo personal y colectivo.

En este sentido, este trabajo se circunscribe en analizar los resultados de dos de las pruebas establecidas por el Ministerio de Educación Nacional (MEN) de Colombia: las condiciones de entrada de los estudiantes antes de comenzar el proceso formativo universitario (desempeño en Saber11) y las condiciones de salida del proceso (desempeño en SaberPRO). Lo anterior enmarca dentro de un sistema de calidad de la educación superior colombiana, donde el concepto de valor agregado no es un simple número sino un punto de partida para el diseño de estrategias que conlleven un mejoramiento de la calidad de la educación impartida.

Con base en lo anterior, la pregunta principal sería: ¿cómo se podría determinar de manera exploratoria el progreso en el rendimiento relativo de los estudiantes antes y después del proceso formativo universitario, tendiendo como base las pruebas asociadas a lectura y escritura, con el ánimo de establecer una propuesta metodológica que sirva para la construcción de un modelo de medición del indicador de calidad Saber11-SaberPRO, por regiones colombianas (Corpes)?

## MARCO TEÓRICO

### Valor agregado

Determinar el grado de influencia positiva que tiene un proceso formativo en los estudiantes colombianos, producto de la contrastación de sus resultados en las pruebas estatales al culminar su secundaria (Saber11), con sus resultados en las pruebas estatales al terminar su educación superior (SaberPRO), es una tarea compleja, si se pretende revisar la adquisición, durante dicha

formación, de habilidades, competencias, etc., que le permitan enfrentar efectivamente los retos reales que se plantean en escenarios futuros.

Diversos autores han realizado estudios relacionados con el tema (Jacob *et al.*, 2009; Ramírez Gómez y Forero Ramírez, 2008 y Medina Durango, 2010), y han identificado que la congruencia entre lo que se enseña y lo que se necesita realmente depende de múltiples y diversos factores, muchos de ellos intangibles, especialmente relacionados con la calidad y con la comunicación efectiva entre el mundo académico-formativo y el mundo laboral. Lastimosamente, estos hallazgos no han sido suficientes, pues algunas instituciones educativas no los han atendido y han eludido incorporar en sus esquemas formativos la sinergia necesaria entre dichos mundos.

Al respecto, Martínez Sánchez (2001) y Vásquez Rizo (2010, p. 70) afirman que no existe una correspondencia entre el sistema educativo y el sistema productivo, es decir, las empresas no reciben de la educación formal las personas idóneas, lo que ha ocasionado que en algunos casos los educandos no se encuentren preparados para afrontar los nuevos retos que para las empresas y sus procesos depara el vertiginoso ritmo del mundo globalizado.

Estudios colombianos apoyados por el MEN y el Instituto Colombiano para la Evaluación de la Educación (Icfes) han mostrado evidente preocupación por provocar dicha sinergia. Ramírez Gómez y Forero Ramírez (2008) han indagado sobre la inserción efectiva de los graduandos universitarios en los mercados laborales y sus ingresos. Por

su parte, Ramírez Gómez *et al.* (2005) se han preocupado por la conformación y consolidación de un Sistema Nacional de Información de la Educación Superior (SNIES) (hoy una realidad); Medina Durango y Posso Suárez (2010) han analizado la relación educación-cambio tecnológico, y Medina Durango (2010) ha indagado sobre los avances y retos de la educación superior, analizando variables como la educación, el empleo, el salario y el valor agregado (VA).

Es por esto que la presente investigación pretende aportar a este análisis, permitiéndole al estudiante que ha cursado nivel medio y superior reconocer sus posibilidades de inclusión, identificar sus competencias estratégicas y realizar procesos de introspección, con la intención de utilizar todo esto como verdaderos elementos significativos y valiosos que le posibiliten alcanzar un grado de distinción y de VA (tanto individual como colectivo), producto de una sinergia cognitiva entre un nivel formativo y otro (proceso lógico y esperado de formación).

El estudiante que pasa por el proceso de formación secundaria-superior necesita convertirse en un ser que aprende e indaga permanentemente, a través de un proceso sistémico y constante; en un ser atento a los cambios repentinos y en un ser siempre dispuesto a adquirir nuevas destrezas, capacidades y competencias, que le permitan apropiarse de la mejor manera los nuevos conocimientos que se adquieren a través de dicho proceso (Shavelson, 2010). Es por esta razón que debe existir, necesariamente, una completa coherencia (si bien no de forma, sí de fondo) entre unas pruebas estatales y otras, para poder generar y determinar un real VA.

En el ámbito educativo, el tema del VA suele volverse un poco más complejo a la hora de medirse en un grupo de personas luego de un proceso, dado que deben analizarse con sumo cuidado, tanto los resultados esperados como las implicaciones que tengan en los sujetos de medición. La valoración del capital humano no ha sido muy estudiada en la literatura, debido a la subjetividad que esto encierra o la disponibilidad de datos indirectos en los que se supone que una determinada inversión en el desarrollo del personal está conectada con un indicador económico de resultado (Amat, 2000, p. 41).

Al respecto, algunos piensan de manera errada que la teoría de la medición se limita solo a aspectos matemáticos, pero se ha establecido que puede ser aplicada a cualquier proceso que implique darle un manejo a los elementos que lo constituyen, luego de haber sido identificados, de acuerdo con su grado de importancia y haberlos interrelacionado. Surge entonces, para este estudio, la necesidad de encontrar una metodología que se aproxime a un estado de medición ideal y a una valoración del capital intelectual adquirido a través de un proceso de formación, en este caso revisando el proceso en dos puntos coyunturales: la culminación de los estudios de nivel medio (Saber11) y la finalización de los estudios de nivel superior (SaberPRO).

En términos educativos, se busca conocer cómo influye el paso de la educación superior en los estudiantes colombianos, es decir, qué tanto ha cambiado el estatus cognitivo del educando al pasar por un nivel superior de formación, pues se analizan los conocimientos adquiridos a través de su formación media y se contrastan con su

“progreso” al culminar la educación superior. En el ámbito de las mediciones educacionales, el VA ha sido considerado como una alternativa más justa, ya que permite comparar con mayor precisión el aporte que hace un establecimiento educacional al aprendizaje de sus estudiantes.

Otros trabajos relacionados con la medición y el VA en contextos educativos se han adelantado en países como Chile (Meckes y Ramírez, 2006), Inglaterra (Ray, 2006) y Estados Unidos (Goldschmidt, 2006), lo cual evidencia que el tema de la medición en dichos escenarios ha sido un tópico de relevancia mundial y sigue siendo un aspecto vigente para la realización de estudios asociados al análisis de impacto educativo (Rockoff, 2004); Shavelson, 2008 y Jacob *et al.*, 2009); situación que evidencia una gran movilización académica en torno al tema.

Shavelson (2008) afirma que el sistema de educación superior es muy complejo como para utilizar solo un tipo de medición, pues se deben medir los resultados educativos en su totalidad para poder identificar verdaderamente las prácticas educativas efectivas e integrar información que sea accesible, comprensible y comparable. Dicha comparación, en el presente trabajo es posible al revisar y contrastar las componentes “Lenguaje” en Saber11 y “Comprensión Lectora” en SaberPRO, pues existe una posibilidad de contrastación adecuada de dichas componentes al pretender evaluar las capacidades y competencias interpretativa, argumentativa y propositiva en el estudiante (Rocha *et al.*, 2011). En este apartado, vale la pena decir que las dos pruebas son independientes, razón por la cual el desempeño en una no condiciona el desempeño en la otra, simplemente, en

este estudio, se toman como referentes de contrastación.

En esta misma dirección, resulta relevante el hecho de medir la comprensión de lectura de los individuos, pues es esta habilidad la que permite la rápida adquisición del conocimiento global y adaptarlo a la realidad, lo cual se relaciona directamente con la productividad (según la estimación de VA realizada en la Encuesta Internacional de Alfabetización de Adultos [IALS], 1998). Otros estudios mencionan que “el VA es medido en términos del grado de comprensión lectora que presentan los egresados de educación universitaria, en relación a lo que obtendrían si sólo hubiesen completado la educación secundaria” (Vargas Cuevas, 2006, p. 134).

A continuación, se presenta una descripción de las pruebas analizadas.

### **Variable de entrada del proceso (Saber11)**

Este examen pretende a) servir como un criterio para el ingreso a la educación superior; b) informar a los estudiantes que aspiran a ingresar a la educación superior acerca de sus competencias en cada una de las áreas evaluadas; c) servir como criterio de autoevaluación y retroalimentación para las instituciones de educación; d) constituirse como base e instrumento para el desarrollo de estudios, y e) servir de criterio para otorgar beneficios educativos y para que algunos estudiantes puedan validar cursos básicos de la educación superior (Rocha *et al.*, 2011).

Esta prueba tiene dos componentes: un núcleo común y otro flexible, dividido en

profundización e interdisciplinar. El núcleo común, obligatorio, evalúa las competencias básicas contextualizadas en las disciplinas de Ciencias Naturales, Ciencias Sociales, Matemáticas, Lenguaje e Idiomas. El núcleo flexible, electivo, evalúa las competencias en niveles de profundización en los contextos disciplinares y en contextos de situaciones y problemáticas que implican la integración de elementos de distintas disciplinas (Rocha *et al.*, 2011).

Para este estudio, el componente “Lenguaje” es la base de formación del universo conceptual que permite al sujeto abstraer la realidad, conceptualizarla, simbolizarla e interpretarla. Es por el lenguaje que el hombre significa la realidad, se constituye como sujeto y crea y recrea la cultura de la cual hace parte (Rocha *et al.*, 2011). Aquí se adopta el análisis del discurso escrito como herramienta de comprensión e interpretación.

### **Variable de salida del proceso (SaberPRO)**

Esta prueba pretende a) comprobar el grado de desarrollo de las competencias de los estudiantes próximos a culminar el pregrado; b) producir indicadores de VA de la educación superior, y c) servir de fuente de información para la construcción de indicadores de evaluación de la calidad de los programas e IES, que fomenten la cualificación de los procesos institucionales y la formulación de políticas en todos los órdenes y componentes del sistema educativo. Esta prueba está conformada, según el Icfes (2010a), por exámenes específicos, para 31 áreas de formación, y exámenes de competencias genéricas (comprensión lectora e inglés obligatorias, lo que facilita la contrastación en este proyecto).



Para este estudio, la “Comprensión Lectora” analiza la capacidad de leer comprensivamente diversos tipos de textos, mediante la aplicación de estrategias comunicativas y lingüísticas específicas y evalúa competencias interpretativas, argumentativas y propositivas (Icfes, 2010a).

A continuación se justifica la realización del estudio por regiones colombianas (Corpes)

### Regiones Corpes

Se decide trabajar por estas regiones, debido a que presentan características homogéneas en algunos aspectos (agrupan departamentos que comparten geografía y características comunes), que permiten desarrollar ejercicios de planificación sectorizados, como el que aquí se propone.

Esta desagregación facilita la labor del investigador y hace mucho más completo y acertado el estudio, pues dichas regiones

requieren análisis separados (tabla 1), ya que existen patrones que aglutinan características propias de cada región, por lo que se puede decir que las disparidades de las regiones de Colombia tienen un referente espacial (Galvis y Meisel Roca, 2010, p. 13), situación que respalda el análisis independiente. Ejemplo de ello es la caracterización que realizan dichos autores, quienes afirman que departamentos como Cundinamarca, Santander (región Oriente) y Antioquia (región Occidente) presentan un nivel de PIB per cápita por encima de la media nacional y se alejan cada vez más de la misma, mientras que departamentos como Atlántico, Bolívar, Magdalena, César, Córdoba, Sucre, La Guajira y San Andrés y Providencia (región costa Caribe) tienden a mantenerse por debajo del promedio nacional. Situación estrechamente relacionada con la cobertura educativa y la calidad de la educación que se imparte (ver pruebas Saber11 y SaberPRO en los últimos años).

**Tabla 1.** Regionalización Corpes en Colombia

CORPES	DEPARTAMENTOS
I. Costa Caribe	1) San Andrés y Providencia, 2) Guajira, 3) Cesar, 4) Magdalena, 5) Atlántico, 6) Bolívar, 7) Sucre, 8) Córdoba
II. Centro oriente	1) Norte de Santander, 2) Santander, 3) Boyacá, 4) Cundinamarca, 5) Tolima, 6) Huila
III. Orinoquía	1) Arauca, 2) Casanare, 3) Vichada, 4) Meta, 5) Guaviare, 6) Guainía, 7) Vaupés
IV. Amazonia	1) Caquetá, 2) Putumayo, 3) Amazonas
V. Occidente	1) Antioquía, 2) Chocó, 3) Caldas, 4) Risaralda, 5) Quindío, 6) Valle, 7) Cauca, 8) Nariño

**Fuente:** Acosta Medina (s.f., p. 2)

Romero (2007), citado por Galvis y Bonilla (2010, p. 16), considera que existen otros factores que igualmente contribuyen a mantener las inequidades. Tal es el caso de las

condiciones del mercado laboral, pues en Colombia hay evidencia de discriminación en dicho mercado, de acuerdo con el grupo étnico o la raza a la cual se pertenezca o la

región de donde se provenga, lo que perjudica el desarrollo homogéneo de la Nación. Por último, Pérez (2007), citado por Galvis y Bonilla (2010, p. 9), encuentra que la ubicación geográfica de la población es importante en términos de la distribución de índices como las necesidades básicas insatisfechas NBI y el índice de calidad de vida, pues hay departamentos y municipios donde el rezago en la calidad de vida está correlacionado con el deterioro encontrado en los departamentos y municipios aledaños.

## METODOLOGÍA

### Análisis de la metodología de valor agregado

Según la OECD (2011), los modelos de VA son una mejora sustancial en comparación con muchas medidas de desempeño escolar. Las comparaciones de puntuaciones brutas de pruebas proporcionan cierta información importante, pero son medidas deficientes del desempeño. El análisis de VA toma en cuenta el rendimiento previo y genera resultados que en gran medida reflejan diferencias en las características contextuales. Los modelos de VA incorporan medidas de aprovechamiento previo y, en algunos casos, características contextuales de los alumnos, lo que permite un análisis más refinado del progreso del desempeño y resulta más eficaz para desentrañar los efectos de diversos factores que afectan su avance. Esto genera mayor precisión al medir el desempeño escolar y más confianza en sus interpretaciones.

Los modelos de VA pueden contribuir a 1) aprender en todos los niveles del sistema educativo al identificar con precisión los aspectos de mayor y menor desempeño;

2) identificar procesos de mejora; 3) establecer sistemas de rendición de cuentas más transparentes y equitativos, que pueden servir para el diseño de incentivos tendientes a elevar el desempeño; 4) desarrollar sistemas de información que permitan a las escuelas analizar y evaluar su desempeño y fortalecer su sistema de evaluación; 5) permitir mayor eficacia y equidad en la asignación de recursos a las áreas críticas, y 6) atender desigualdades socioeconómicas, que puedan estar ocultas en el ámbito escolar mediante medidas de desempeño indiscriminadas, imprecisas e inequitativas.

Es así como se propone un indicador para medir el VA por regiones Corpes: el porcentaje de efectividad máxima, que se presenta a continuación.

### Porcentaje de efectividad máxima

En primera instancia, se pretende mostrar los desempeños estudiantiles en contexto. Esto resalta la importancia de poder llevar todo a un mismo referente que coadyuve a la identificación de elementos de comparabilidad, pero lógicamente teniendo en cuenta las variaciones y distribuciones asociadas a cada variable objeto de análisis. Es una propuesta metodológica que, como toda propuesta académica, se espera que se encuentre en constante discusión y ajuste para posteriores investigaciones en la temática.

Esto significa que no bastaría con los rendimientos absolutos, puesto que las condiciones de la población marcan una tendencia importante, tendencia sobre la cual se deben analizar los desempeños individuales. Dado lo anterior, se usarán los porcentajes de efectividad máxima como muestra

relativa de los puntajes. El término porcentaje de efectividad máxima es una construcción terminológica (propuesta), en la cual se intentan describir los puntajes obtenidos por los estudiantes en función del máximo alcanzable para cada cohorte de evaluación. Desde esta perspectiva, es una razón que tiene en cuenta en el numerador el puntaje promedio alcanzado por cada estudiante en la prueba respectiva y en el denominador el puntaje máximo que podría obtener cualquier estudiante, considerando los promedios nacionales para la fecha en la cual el estudiante presentó la prueba (análisis por cohortes), más tres veces la desviación estándar de la prueba en la misma fecha (análisis por cohortes).

Dicha razón se ampara en la regla empírica de la distribución normal (Lind, Marchal y Watchen, 2008), en la cual la gran mayoría de los desempeños se encontrarán entre el promedio y más o menos tres desviaciones estándar, por tanto, una aproximación a los máximos valores estaría concentrada en analizar el promedio más las tres desviaciones estándar de cada cohorte o periodo en la cual se presentaron las pruebas censales.

Por este motivo, se identifican las regiones Corpes por analizar, y se toman los puntajes obtenidos por los estudiantes en “Comprensión Lectora” (SaberPRO). Estos puntajes se llevan a una expresión relativa, como lo es el porcentaje de efectividad máxima. En este caso, se tendrían los desempeños en el contexto del total de estudiantes que presentaron dicha prueba en el país. La otra medición de este mismo individuo (muestra pareada) se obtiene a partir de una visión retrospectiva, donde se busca para cada estudiante su desempeño en “Lenguaje”

(Saber11). En este procedimiento, también se expresan los rendimientos de una manera relativa, a través del porcentaje de efectividad máxima. Se muestran las diferencias entre los porcentajes de efectividad máxima antes y después del proceso formativo, esperando que exista una mejora en el posicionamiento relativo de los estudiantes, luego de tener una formación universitaria.

En los resultados se describe la proporción de estudiantes que igualan o superan su porcentaje de efectividad máxima (superan sus desempeños relativos), de la misma manera que su complemento evidencia aquellos que disminuyeron sus posiciones relativas. La proporción de estudiantes que igualan o superan marca un indicador positivo en la medida en que se aproxime al 100 %, situación que permitirá empíricamente mostrar una incidencia del proceso formativo.

La proporción de estudiantes que igualan o superan su porcentaje de efectividad máxima (SaberPRO-Saber11) se define de acuerdo con la siguiente notación:

$$P = \frac{\sum pemsaberpro \geq pemsaber11}{n}$$

La sumatoria permite contabilizar el número de estudiantes cuyos porcentajes de efectividad máxima en SaberPRO son mejores o iguales que los porcentajes de efectividad máxima en Saber11, donde: *pemsaberpo* = porcentaje de efectividad máxima en SaberPRO; *pemsaber 11* = porcentaje de efectividad máxima en Saber11, y *n* = número total de estudiantes de la región Corpes analizada que presentaron SaberPRO.

Por otra parte, también se emplea un indicador utilizado con frecuencia: promedio del porcentaje de efectividad máxima en SaberPRO de la cohorte de salida/promedio del porcentaje de efectividad máxima en Saber11 de cohorte de salida al momento de ingreso al programa. En esta medición, una relación mayor que 1 indicará que el programa ayudó a elevar la posición relativa del grupo frente al promedio nacional y, por tanto, agregó valor. Una relación por debajo de la unidad indica lo contrario. La notación del indicador se establece de la siguiente manera:

$$I = \frac{\frac{\sum_{i=1}^n pemsaberpro_i}{n}}{\frac{\sum_{i=1}^n pemsaber11_i}{n}}$$

En la metodología propuesta se continúan explorando evidencias que lleven a pensar en el progreso relativo de los estudiantes y de la incidencia del proceso formativo en sus posiciones relativas. Por tanto, se plantean tres comparaciones estadísticas donde se desea conocer si existen progresos significativos, acordes con la distribución de los puntajes, entre las puntuaciones antes y después de la intervención educativa. Para ello, se emplea la prueba para muestras pareadas: prueba t (supuesto distribución normal), prueba de Wilcoxon y prueba de signos (sin supuesto sobre la distribución).

### Prueba dependientes o pareadas

En algunas oportunidades se tienen muestras dependientes o “pareadas”, ya sea

porque se realizaron dos observaciones sobre el mismo elemento o individuo, o porque se hizo un pareo de elementos de acuerdo con ciertas características. En estos casos lo que se tiene, realmente, es una muestra de “pares” de elementos y lo que interesa es determinar si, en promedio, las diferencias entre esos pares difieren significativamente o no de cero.

### Prueba t para muestras relacionadas

Tomando como referencia SPSS. Guía para el análisis de datos (Universidad de Cádiz), se tiene que la prueba t para dos muestras relacionadas permite contrastar hipótesis referidas a la diferencia entre dos medias relacionadas. Ahora se dispone de una población de diferencias con media  $\mu_D$ , obtenida al restar las puntuaciones del mismo grupo de casos en dos variables diferentes o en la misma variable medida en dos momentos diferentes (de ahí que se hable de muestras relacionadas). De esa población de diferencias se extrae una muestra aleatoria de tamaño  $n$  y se emplea la media  $\bar{Y}_D$  de esas  $n$  diferencias para contrastar la hipótesis de que la media  $\mu_D$  de la población de diferencias vale cero.

Desde la estadística, este contraste es idéntico al de prueba t para una muestra. La única diferencia es que ahí se tenía una muestra de puntuaciones obtenida al medir una sola variable y ahora se cuenta con dos muestras relacionadas (o una muestra de pares de puntuaciones), que se convierten en una sola muestra de diferencias restando las puntuaciones de cada par. El estadístico o prueba T sigue siendo una tipificación de la media muestral de las diferencias  $\bar{Y}_D$ :

$$T = \frac{\bar{Y}_D - \mu_D}{\hat{\sigma}_{\bar{Y}_D}} = \frac{\bar{Y}_D - \mu_D}{S_D / \sqrt{n}}$$

( $S_D$  se refiere a la desviación típica insesgada de las  $n$  diferencias).

Este estadístico  $t$  se distribuye según el modelo  $t$  de student con  $n-1$  grados de libertad y, por tanto, permite conocer la probabilidad asociada a los diferentes valores  $\bar{Y}_D$  que es posible obtener en muestras aleatorias de tamaño  $n$ .

Al igual que antes, para que el valor  $t$  se ajuste apropiadamente al modelo de distribución de probabilidad  $t$  de student, es necesario que la población de diferencias sea normal. No obstante, con tamaños muestrales grandes el ajuste del estadístico  $t$  a la distribución  $t$  de student es lo suficientemente bueno, incluso con poblaciones originales alejadas de la normalidad.

### Prueba de Wilcoxon

Tomando como referencia SPSS. Guía para el análisis de datos (Universidad de Cádiz) se tiene que esta es una prueba utilizada en el análisis no paramétrico, que permite analizar datos de diseños con medidas repetidas. Se supone tomar dos medidas ( $X_i$  e  $Y_i$ ) a un grupo de  $m$  sujetos y se calculan las diferencias en valor absoluto entre las dos puntuaciones de cada par:  $D_i = |X_i - Y_i|$ ; ( $i = 1, 2, \dots, m$ ).

Se descartan las  $D_i$  nulas y se considerará únicamente las  $n$  diferencias  $D_i$  no nulas ( $n - m$ ). Se asignan rangos ( $R_i$ ) desde 1 hasta  $n$  a esas  $D_i$  no nulas: el rango 1 a la  $D_i$  más pequeña, el rango 2 a la  $D_i$  más pequeña de las restantes, el rango  $n$  a la  $D_i$

más grande (si existen empates, se resuelven asignando el promedio de los rangos). Se suman, por un lado, los  $R_i^+$ , es decir, los rangos correspondientes a las  $D_i$  con  $X_i > Y_i$ , y se le llama  $S_+$  a esta suma; y se suman, por otro lado, los  $R_i^-$ , es decir, los rangos correspondientes a las  $D_i$  con  $X_i < Y_i$ , y se le llama  $S_-$  a esta otra suma. Si se supone que las puntuaciones  $X_i$  e  $Y_i$  proceden de poblaciones con la misma mediana ( $Mdn_x = Mdn_y$ ), se debe esperar que:  $P(X_i < Y_i) = P(X_i > Y_i)$ .

Por lo que, si la hipótesis  $H_o: Mdn_x = Mdn_y$  es verdadera, en una muestra aleatoria de  $n$  observaciones, se deben encontrar aproximadamente tantos valores  $X_i > Y_i$  como valores  $X_i < Y_i$  (salvando, por supuesto, las fluctuaciones atribuibles al azar muestral). Pero si la distribución de las diferencias es simétrica (lo cual exige escala de intervalo o razón), las  $D_i$  positivas se alejarán de cero en igual medida que las  $D_i$  negativas; de donde es fácil deducir que:

$S_+ = \sum R_i^+ \approx S_- = \sum R_i^-$ . Es decir, si  $Mdn_x = Mdn_y$ , la distribución de las diferencias  $D_i$  es simétrica,  $S_+$  y  $S_-$  tomarán valores parecidos. Por tanto, una fuerte discrepancia entre  $S_+$  y  $S_-$  hará dudar de la veracidad de  $H_o$ . De modo que se procede a utilizar los valores  $S_+$  y  $S_-$  para obtener información sobre la hipótesis  $H_o: Mdn_x = Mdn_y$  (Wilcoxon, 1945, 1949).

Con tamaños muestrales pequeños no resulta complicado obtener la distribución exacta del estadístico  $S_+$  (o  $S_-$ ), pero es más rápido obtener una tipificación de  $S$  ( $S$  se refiere al menor de  $S_+$  y  $S_-$ ), cuya distribución se aproxima, conforme al tamaño muestral va aumentando, al modelo de probabilidad normal  $N(0, 1)$ :

$$Z = \frac{S - \frac{n(n+1)}{4}}{\sqrt{\frac{n(n+1)(2n+1)}{24} - \sum_{i=1}^k \frac{t_i^3 - t_i}{48}}} \quad X_i < Y_i$$

Donde ( $k$  se refiere al número de rangos distintos en los que existen empates y  $t_i$ , al número de puntuaciones empatadas en el rango  $i$ ). El SPSS ofrece el nivel crítico bilateral resultante de multiplicar por 2 la probabilidad de obtener valores menores o iguales que  $Z$ .

### Prueba de los signos

Tomando como referencia SPSS. Guía para el análisis de datos (Universidad de Cádiz), se describe esta prueba como una herramienta utilizada en el análisis no paramétrico, que permite analizar datos de diseños con medidas repetidas. Al igual que la prueba de Wilcoxon, esta prueba permite contrastar la hipótesis de igualdad entre dos medianas poblacionales. Pero, mientras la prueba de Wilcoxon aprovecha la información ordinal de los datos (aunque exige nivel de medida de intervalo o razón), la prueba de los signos solo aprovecha de los datos sus propiedades nominales (aunque exige nivel de medida al menos ordinal). Se supone que tomando dos medidas ( $X_i$  e  $Y_i$ ) a un grupo de  $m$  sujetos y calculando las diferencias en valor absoluto entre las dos puntuaciones de cada par, resulta:  $D_i = |X_i - Y_i|$ ; ( $i = 1, 2, \dots, m$ ).

Se desechan las  $D_i$  nulas y se consideran únicamente las  $n$  diferencias  $D_i$  no nulas ( $n \leq m$ ). Si se supone que las puntuaciones  $X_i$  e  $Y_i$  proceden de poblaciones con la misma mediana ( $Mdn_x = Mdn_y$ ), debe verificarse que  $P(X_i < Y_i) = P(X_i > Y_i) = 0.5$ . De modo que, si la hipótesis  $H_o: Mdn_x = Mdn_y$  es verdadera, en una

muestra aleatoria den observaciones se deben encontrar aproximadamente tantos valores  $X_i > Y_i$  como valores  $X_i < Y_i$ , es decir, aproximadamente tantas diferencias  $D_i$  positivas como negativas (salvando, por supuesto, las fluctuaciones atribuibles al azar propio del proceso de muestreo). Desde estas condiciones, las variables  $n_+$  = número de signos positivos y  $n_-$  = número de signos negativos se distribuyen según el modelo binomial con parámetros  $n$  y  $\pi = 0,50$ . De modo que se puede utilizar la distribución binomial para conocer las probabilidades asociadas a  $n_+$  y  $n_-$ , y, basándose en ellas, contrastar la hipótesis  $H_o: Mdn_x = Mdn_y$ .

Si  $n \leq 25$ , el SPSS toma el valor  $k = \min(n_+, n_-)$  y utilizando las probabilidades de la distribución binomial calcula el nivel crítico bilateral resultante de multiplicar por 2 la probabilidad de obtener valores iguales o menores que  $k$ . Si  $n > 25$ , el SPSS tipifica el valor de  $k$  (utilizando corrección por continuidad) y ofrece el nivel crítico resultante de multiplicar por 2 la probabilidad de encontrar valores iguales o menores que

$$Z: Z = \frac{k + 0,5 - n/2}{0,5\sqrt{n}}$$

### Fuentes de información

Las fuentes de información que se emplearán en este proyecto serán las bases de datos de las pruebas Saber11 y SaberPRO, aplicadas en el periodo 2000-2009. Se tomará una muestra y se aparearán los datos de los estudiantes por las variables "documento de identidad", "nombre", "institución educativa", "departamento y región Corpes". Se debe tener presente que los datos deben permitir la trazabilidad del estudiante en el periodo de estudio.

## RESULTADOS

Numerosos investigadores han abordado el tema de comparar diferentes modelos para separar y medir los efectos de la escuela, de aquellas características de entrada de los estudiantes, tales como rendimiento previo y factores socioeconómicos (Sammons, Thomas y Mortimore, 1997). Según Froemel Andrade (2003, p. 168),

los hallazgos de Thomas y Mortimore (1996) y los de Sammons *et al.* (1994) señalan que un número de escuelas (quizá un 20 % del total) pueden obtener diferentes resultados cuando se cuenta para el análisis con datos disponibles respecto de las condiciones de entrada, (en contraste) con los resultados basados solamente en datos del contexto de los estudiantes —tales como: edad, género, etc.

Más aún, Thomas y Mortimore (1996) y Sammons *et al.* (1994) plantean que: “es interesante constatar que cuando se han usado adecuadas ‘líneas base’, la inclusión de información socioeconómica en los cálculos contribuye poco, en cuanto a explicar las diferencias entre los estudiantes”.

Dado lo anterior, el presente estudio involucra ahora un análisis a través del tiempo de los rendimientos obtenidos por un mismo individuo antes y después de la intervención académica. La línea base la constituyen los rendimientos en “Lenguaje” en Saber11 y los resultados se observan a través de “Comprensión Lectora” en SaberPRO. Como manifiestan Sammons, Thomas y Mortimore (1997), citados por Froemel Andrade (2003, p. 170): “es esencial que, donde sea posible, los sistemas de ‘valor agregado’

estén basados en un diseño longitudinal para medir el progreso de los estudiantes”.

### **Cálculo del porcentaje de aproximación al referente de calidad (efectividad máxima)**

Este apartado tiene el propósito fundamental de llevar las dos medidas que obtiene un estudiante (Saber11 y SaberPRO) a un mismo lenguaje, por medio del cual se pueda interactuar y conocer de qué manera se han experimentado progresos o decrecimientos. Desde esta perspectiva, para “Lenguaje”, en Saber11, se espera una distribución normal, toda vez que es un proceso de medición educativa, además de las características censales de la prueba. A través de la regla empírica de la probabilidad se conoce que entre los valores de la media  $\pm$  tres desviaciones estándar, se encuentra el 99,7 % de las observaciones, en este caso los puntajes obtenidos en la componente de “s” pruebas de “Lenguaje”.

Se toma entonces para este estudio, como una medida de mayor efectividad máxima, aquella puntuación que alcance a ubicarse en el promedio nacional más tres desviaciones estándar, en “Lenguaje”, del año en el que el individuo presentó la prueba Saber11. Tómese, para ilustración, al “Individuo 1”, cuyo puntaje en “Lenguaje” fue 58. Este individuo presentó el examen en el segundo periodo de 2001, ocasión en la cual dicha prueba alcanzó un promedio nacional de 46,51 y una desviación estándar de 6,02. El porcentaje de efectividad máxima para dicho individuo está dado por  $58 / (46,51 + 3 \times 6,02) = 0,8982$ . Análogamente, se procede para “Comprensión Lectora” en SaberPRO, teniendo en cuenta los



elementos o referentes de análisis anteriormente mencionados.

De esta manera, se llega a dos medidas %efecSaber11 (porcentaje de efectividad máxima en “Lenguaje” de Saber11) y %efecSaberPRO (porcentaje de efectividad máxima en “Comprensión Lectora” de SaberPRO). Dichas medidas tienen en cuenta las características intrínsecas a sus respectivas distribuciones y, de esta manera, al ser expresadas como porcentajes de efectividad pueden ser fácilmente comparables en términos del progreso evidenciado antes y después del proceso de intervención educativa.

A continuación, se hace efectiva la ecuación contenida en la metodología para el cálculo del porcentaje de aproximación al referente de calidad (efectividad máxima), para cada Corpes:

En la región **Amazonia**:

$$P = \frac{\sum pemsaberpro \geq pemsaber11}{n} = \frac{924}{1164} = 79,4\%$$

En la región **Centro Oriente**:

$$P = \frac{\sum pemsaberpro \geq pemsaber11}{n} = \frac{54701}{90197} = 60,6\%$$

En la región **Costa Atlántica**:

$$P = \frac{\sum pemsaberpro \geq pemsaber11}{n} = \frac{18706}{28080} = 66,6\%$$

En la región **Occidente**:

$$P = \frac{\sum pemsaberpro \geq pemsaber11}{n} = \frac{33865}{52229} = 64,8\%$$

En la región **Orinoquía**:

$$P = \frac{\sum pemsaberpro \geq pemsaber11}{n} = \frac{1709}{2575} = 66,4\%$$

Como referente inicial y de tipo descriptivo, se puede apreciar, en la totalidad de las regiones Corpes, que el porcentaje de individuos que supera su porcentaje de efectividad con respecto al ingreso se encuentra por encima del 60 % (mejoraron su posición luego del proceso de intervención). Se observan mayores progresos del indicador en regiones como la Amazonia, debido, en gran parte, a las condiciones de arranque o línea base del indicador.

Aunque se presentan inicialmente indicadores satisfactorios, vale la pena resaltar el análisis del complemento de la proporción: aquellos estudiantes que no están superando sus desempeños con respecto al estándar de calidad elegido. En este sentido, un frente de trabajo futuro, y que entra a ser parte de la discusión, debería enfocarse en identificar o caracterizar dichas poblaciones vulnerables, de tal manera que los procesos de intervención ejerzan un liderazgo propositivo en las nuevas cohortes de estudiantes que ingresan a estos planteles educativos y que guardan características comunes, siendo susceptibles de decrecimientos en sus niveles de efectividad con respecto al estándar.

Por otro lado, en cuanto al indicador relacionado con el promedio del porcentaje de efectividad máxima en la prueba SaberPRO de la cohorte de salida/promedio del porcentaje de efectividad máxima en la prueba Saber11 de cohorte de salida al momento de ingreso al programa, la ecuación también se hace efectiva para cada región Corpes:



En la región **Amazonia**:

$$I = \frac{\frac{\sum_{i=1}^n pemsaberpro_i}{n}}{\frac{\sum_{i=1}^n pemsaber11_i}{n}} = \frac{0,7964}{0,7132} = 1,12$$

En la región **Centro Oriente**:

$$I = \frac{\frac{\sum_{i=1}^n pemsaberpro_i}{n}}{\frac{\sum_{i=1}^n pemsaber11_i}{n}} = \frac{0,7993}{0,7706} = 1,04$$

En la región **Costa Atlántica**:

$$I = \frac{\frac{\sum_{i=1}^n pemsaberpro_i}{n}}{\frac{\sum_{i=1}^n pemsaber11_i}{n}} = \frac{0,7915}{0,7454} = 1,06$$

En la región **Occidente**:

$$I = \frac{\frac{\sum_{i=1}^n pemsaberpro_i}{n}}{\frac{\sum_{i=1}^n pemsaber11_i}{n}} = \frac{0,8007}{0,7586} = 1,06$$

En la región **Orinoquía**:

$$I = \frac{\frac{\sum_{i=1}^n pemsaberpro_i}{n}}{\frac{\sum_{i=1}^n pemsaber11_i}{n}} = \frac{0,7954}{0,7489} = 1,06$$

En este caso también se observa una mejora en el valor del porcentaje de efectividad

máxima antes y después de la intervención educativa, lo que a su vez se constituye en un indicador que refuerza el indicador anterior, en el sentido de afirmar, desde un punto de vista exploratorio inicial, que sí se ha dado un impacto positivo en los procesos formativos a través del mejoramiento del desempeño relativo de los estudiantes.

### Comparación de los resultados con metodologías tradicionales empleadas en educación

Este apartado del estudio hace referencia a metodologías tradicionales, con el fin de observar elementos de contrastación teórica que coadyuvan a la explicación y justificación de los indicadores propuestos.

El porcentaje de efectividad máxima es una propuesta que intenta tomar en consideración una construcción teórica con respecto al ideal máximo alcanzable por un estudiante que toma una prueba en un año particular. Este ideal se ampara en las definiciones alrededor de la regla empírica de probabilidad, que puede servir para la discusión y construcción en colectivo de nuevas miradas hacia los estudios provenientes de la rama de investigación educativa.

Como propuesta que constituye, es necesario que este indicador sea contrastado a la luz de los percentiles, dado que, se bien es cierto que existe un mejoramiento en cuanto a los referentes de calidad a lo largo de las dos pruebas censales objeto de análisis, también es de vital importancia analizar el comportamiento en general de los estudiantes con respecto a la cohorte que presenta el examen. En otras palabras, puede que haya

existido un mejoramiento significativo en los estudiantes, en cuanto a los niveles de calidad deseables, pero también se debe conocer si en términos generales la situación también ha presentado un mejoramiento.

Aquí vale la pena mencionar que un decremento en el percentil a nivel agregado no necesariamente indica un descenso en la calidad de un estudiante. Debe analizarse con sumo cuidado y teniendo en cuenta múltiples factores, dado que una prueba puede ser calibrada o diseñada con ciertos criterios de potencia que posibilitan el hecho de que más o menos miembros de la población puedan concentrarse en umbrales diferentes; es decir, no necesariamente se debe a la calidad de un estudiante al presentar la prueba sino también a condiciones intrínsecas del instrumento de medición; así como también se debe a la cualificación de la población, lo que puede llevar a aumentos de umbrales definidos para los percentiles.

Para llevar a cabo la asignación de percentiles a un determinado puntaje, se realiza la estandarización de cada puntaje a un coeficiente tipificado en la normal estándar a través del promedio y desviación estándar generales de cada prueba, en cada año específico de análisis y posteriormente, se halla el porcentaje correspondiente en la curva normal. En este caso se estaría obteniendo la aproximación de los percentiles para cada puntaje.

Se ha decidido adoptar como forma de cálculo (con interpretaciones obviamente diferentes), una notación similar a la utilizada en el momento de establecer los porcentajes de efectividad máxima por región Corpes.

En la región **Amazonia**:

$$P = \frac{\sum \text{percsaberpro} \geq \text{percsaber11}}{n} = \frac{455}{1164} = 39,1\%$$

$$I = \frac{\frac{\sum \text{percsaberpro}_i}{n}}{\frac{\sum \text{percsaber11}_i}{n}} = \frac{0,5094}{0,6951} = 0,73$$

Aquellos estudiantes que obtienen un percentil superior en SaberPRO con respecto a Saber11 constituyen un porcentaje del 40 %, es decir, que estos estudiantes mejoran su posicionamiento de acuerdo con el general de estudiantes. Con fines simplemente descriptivos, se puede apreciar que el percentil promedio en las pruebas Saber11 es de 0,6951 mientras que este mismo indicador en SaberPRO constituye 0,5094. Esta información, si bien hace parte de un análisis adicional, no es antagónica sino complementaria al denominado porcentaje de efectividad, dado que con respecto al examen los estudiantes alcanzan mayores niveles de efectividad, pero al ser un buen porcentaje de ellos cada vez el mejoramiento del conglomerado de los percentiles, es más difícil, como se alcanza a apreciar en estos indicadores.

En los estudiantes se pueden apreciar mejores niveles de abstracción y desarrollo de habilidades que les permitan superar las pruebas, pero esto se presenta en un gran número de ellos. En futuros trabajos, valdría la pena estudiar el diseño de las pruebas como elemento condicionante en la fluctuación, tanto de relación con el ideal alcanzable como de los percentiles propiamente dichos.

A continuación se presentan los resultados para el resto de regiones Corpes, que plantea

en todos los cuatro casos la misma problemática con similares interpretaciones a las ya expuestas para la región de la Amazonia.

En la región **Centro Oriente**:

$$P = \frac{\sum_{i=1}^n \text{percsaberpro}_i \geq \text{percsaber11}}{n} = \frac{25510}{90197} = 28,3\%$$

$$I = \frac{\frac{\sum_{i=1}^n \text{percsaberpro}_i}{n}}{\frac{\sum_{i=1}^n \text{percsaber11}_i}{n}} = \frac{0,5129}{0,6948} = 0,74$$

En la región **Costa Atlántica**:

$$P = \frac{\sum_{i=1}^n \text{percsaberpro}_i \geq \text{percsaber11}}{n} = \frac{8912}{28080} = 31,7\%$$

$$I = \frac{\frac{\sum_{i=1}^n \text{percsaberpro}_i}{n}}{\frac{\sum_{i=1}^n \text{percsaber11}_i}{n}} = \frac{0,4779}{0,6320} = 0,76$$

En la región **Occidente**:

$$P = \frac{\sum_{i=1}^n \text{percsaberpro}_i \geq \text{percsaber11}}{n} = \frac{16895}{52229} = 32,3\%$$

$$I = \frac{\frac{\sum_{i=1}^n \text{percsaberpro}_i}{n}}{\frac{\sum_{i=1}^n \text{percsaber11}_i}{n}} = \frac{0,5218}{0,6687} = 0,78$$

En la región **Orinoquía**:

$$P = \frac{\sum_{i=1}^n \text{percsaberpro}_i \geq \text{percsaber11}}{n} = \frac{841}{2575} = 32,7\%$$

$$I = \frac{\frac{\sum_{i=1}^n \text{percsaberpro}_i}{n}}{\frac{\sum_{i=1}^n \text{percsaber11}_i}{n}} = \frac{0,4975}{0,6432} = 0,77$$

## Pruebas pareadas

Las muestras pareadas son muy útiles cuando se desean hacer comparaciones, porque al realizar las mediciones en los mismos individuos mantienen muchos factores constantes y facilitan la evaluación de los efectos que se quieren analizar. Estas muestras aparecen con cierta frecuencia en la investigación educativa, cuando se desean determinar los efectos de un método de enseñanza en un mismo grupo de alumnos o cuando se desean comparar pruebas iniciales de diagnóstico con pruebas finales. En estas situaciones se debe estar muy atento para no caer en el error de analizar los resultados como si fueran muestras independientes.

A continuación, se presentan los principales resultados derivados de la prueba *t*, la prueba de Wilcoxon y la prueba de signos para cada una de las regiones Corpes analizadas, con el fin de conocer el impacto en el proceso educativo de manera global.

### La prueba *t*

La tabla 2 recoge, para cada variable, la media, el número de casos, la desviación típica y el error típico de la media. Se observan los porcentajes de efectividad máxima promedio para cada prueba relacionada con “Comprensión Lectora” antes y después del proceso educativo. Cuando se toma la referencia del coeficiente de variación (relación entre la desviación estándar y el promedio) se reconoce que la distribución de los porcentajes para ambos casos es de naturaleza homogénea; situación consistente con lo planteado a lo largo del documento. En esta tabla, es importante señalar que la región Occidente presenta en promedio los mayores porcentajes de efectividad en “Comprensión Lectora”.

**Tabla 2.** Media, número de casos, desviación típica, error típico de la media y porcentajes de efectividad máxima promedio por regiones Corpes para las componentes “Comprensión Lectora” y “Lenguaje”.

Paired Samples Statistics						
CORPES			Mean	N	Std. Deviation	Std. Error Mean
AMAZONIA	Pair 1	Efectividad Comprensión Lectora	,7964	1164	,06796	,00199
		Efectividad Lenguaje	,7132	1164	,09804	,00287
CENTRO ORIENTE	Pair 1	Efectividad Comprensión Lectora	,7993	90197	,07118	,00024
		Efectividad Lenguaje	,7706	90197	,10353	,00034
COSTA ATLÁNTICA	Pair 1	Efectividad Comprensión Lectora	,7915	28080	,07190	,00043
		Efectividad Lenguaje	,7454	28080	,10452	,00062
OCCIDENTE	Pair 1	Efectividad Comprensión Lectora	,8007	52229	,07168	,00031
		Efectividad Lenguaje	,7586	52229	,10378	,00045
ORINOQUÍA	Pair 1	Efectividad Comprensión Lectora	,7954	2575	,07242	,00143
		Efectividad Lenguaje	,7489	2575	,10076	,00199

**Fuente:** elaboración propia, utilizando la base de datos (2010b) y el repositorio digital del Icfes (2010c)

**Tabla 3.** Coeficiente de correlación de Pearson y nivel crítico bilateral que corresponde bajo la hipótesis de independencia por regiones Corpes para las componentes “Comprensión Lectora” y “Lenguaje”.

Paired Samples Correlations					
CORPES			N	Correlation	Sig.
AMAZONIA	Pair 1	Efectividad Comprensión Lectora & Efectividad Lenguaje	1164	,205	,000
CENTRO ORIENTE	Pair 1	Efectividad Comprensión Lectora & Efectividad Lenguaje	90197	,237	,000
COSTA ATLÁNTICA	Pair 1	Efectividad Comprensión Lectora & Efectividad Lenguaje	28080	,221	,000
OCCIDENTE	Pair 1	Efectividad Comprensión Lectora & Efectividad Lenguaje	52229	,213	,000
ORINOQUÍA	Pair 1	Efectividad Comprensión Lectora & Efectividad Lenguaje	2575	,238	,000

**Fuente:** elaboración propia, utilizando la base de datos (2010b) y el repositorio digital del Icfes (2010c)

La tabla 3 muestra el coeficiente de correlación de Pearson entre ambas variables, junto con el nivel crítico bilateral que le corresponde bajo la hipótesis de independencia. Se observa que los valores están por debajo de 0,05 en la significancia, lo que indica que es estadísticamente significativa la correlación no cero (es decir existe asociación), con un 95 % de nivel de confiabilidad. En este caso se aprecia una asociación significativa entre el *%efecSaber11* y su posterior

*%efecSaberPRO*; se resalta además que no existe una correlación muy grande entre los dos desempeños, situación interesante en la reflexión del impacto, dado que no necesariamente los estudiantes que obtuvieron menores desempeños relativos en Saber11 estarían “destinados” a obtener mayores desempeños relativos en SaberPRO. Lo anterior, desde una perspectiva proactiva, afianzaría la presencia de los procesos de intervención educativa llevados a cabo.

**Tabla 4.** Media, desviación típica, error típico de la media, intervalo de confianza para la diferencia entre el *%efecSaberPRO* y el *%efecSaber11*, valor del estadístico t, grados de libertad y nivel crítico bilateral por regiones Corpes para las componentes “Comprensión Lectora” y “Lenguaje”

Paired Samples Test										
CORPES			Paired Differences					t	df	Sig. (2-tailed)
			Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
						Lower	Upper			
AMAZONIA	Pair 1	Efectividad Comprensión Lectora - Efectividad Lenguaje	,08321	,10720	,00314	,07705	,08938	26,482	1163	,000
CENTRO ORIENTE	Pair 1	Efectividad Comprensión Lectora - Efectividad Lenguaje	,02876	,11089	,00037	,02804	,02949	77,897	90196	,000
COSTA ATLÁNTICA	Pair 1	Efectividad Comprensión Lectora - Efectividad Lenguaje	,04612	,11300	,00067	,04480	,04744	68,398	28079	,000
OCCIDENTE	Pair 1	Efectividad Comprensión Lectora - Efectividad Lenguaje	,04211	,11285	,00049	,04115	,04308	85,287	52228	,000
ORINOQUÍA	Pair 1	Efectividad Comprensión Lectora - Efectividad Lenguaje	,04649	,10921	,00215	,04227	,05071	21,602	2574	,000

**Fuente:** elaboración propia, utilizando la base de datos (2010b) y el repositorio digital del Icfes (2010c).

En la tabla 4, en la primera mitad, se muestran tres estadísticos referidos a las diferencias entre cada par de puntuaciones: la media, la desviación típica y el error típico de la media. La siguiente columna contiene el intervalo de confianza para la diferencia entre el *%efecSaberPRO* y el *%efecSaber11*. Se puede estimar, por tanto, con una confianza del 95 %, por la verdadera diferencia entre las medias, que por ejemplo en el caso de la región Amazonia, este se encuentra entre 0,07705 y 0,08938. La segunda mitad de la tabla informa sobre el valor del

estadístico t, sus grados de libertad (gl) y el nivel crítico bilateral (sig.bilateral). Aquí se puede observar que el valor del nivel crítico es pequeño (0,000) y, aunado esto al hecho de que el intervalo de las diferencias siempre es positivo para todas las regiones (Amazonia, Centro Oriente, Costa Atlántica, Occidente y Orinoquía), se puede rechazar la hipótesis de igualdad de medias y, por tanto, se concluye que el porcentaje de efectividad máxima (SaberPRO) medio es significativamente mayor que el porcentaje de efectividad máxima (Saber11) medio,

lo cual muestra una influencia positiva del proceso educativo adelantado.

### La prueba Wilcoxon

La tabla 5 ofrece algunos estadísticos descriptivos para las dos variables seleccionadas: el número de casos válidos en ambas variables, la media, la desviación típica, el valor más pequeño, el más grande y los cuartiles. Se puede apreciar que el mínimo porcentaje de efectividad máxima alcanzado por un estudiante se obtuvo en “Lenguaje” (0,17) de Saber11, tanto en la región Occidente como en la región Centro Oriente.

En función del cuartil 2 (mediana), también se evidencia una situación para la Amazonia, en la cual la mitad de los estudiantes obtiene un valor de porcentaje de efectividad inferior a 0,7120 en “Lenguaje” de Saber11, mientras que en la región Occidente esta misma medida se encuentra en 0,76; de esta manera, se exponen niveles apropiados de desarrollo educativo. Como se puede apreciar, las medianas de los porcentajes de efectividad en “Comprensión Lectora” de SaberPRO son ligeramente mayores a las obtenidas en “Lenguaje” de Saber11, lo que reafirma el mejoramiento con respecto a referentes de calidad estatales.

**Tabla 5.** Número de casos válidos en ambas variables, media, desviación típica, valor más pequeño, valor más grande y cuartiles por regiones Corpes para las componentes “Comprensión Lectora” y “Lenguaje”

CORPES		N	Mean	Std. Deviation	Minimum	Maximum	Percentiles		
							25th	50th (Median)	75th
AMAZONIA	Efectividad Comprensión Lectora	1164	,7964	,06796	,48	1,08	,7586	,7984	,8402
	Efectividad Lenguaje	1164	,7132	,09804	,37	1,12	,6438	,7120	,7744
CENTRO ORIENTE	Efectividad Comprensión Lectora	90197	,7993	,07118	,40	1,20	,7573	,8038	,8460
	Efectividad Lenguaje	90197	,7706	,10353	,17	1,32	,7058	,7729	,8363
COSTA ATLÁNTICA	Efectividad Comprensión Lectora	28080	,7915	,07190	,45	1,17	,7504	,7957	,8386
	Efectividad Lenguaje	28080	,7454	,10452	,29	1,41	,6737	,7434	,8160
OCCIDENTE	Efectividad Comprensión Lectora	52229	,8007	,07168	,44	1,17	,7579	,8061	,8482
	Efectividad Lenguaje	52229	,7586	,10378	,17	1,32	,6879	,7594	,8254
ORINOQUÍA	Efectividad Comprensión Lectora	2575	,7954	,07242	,44	1,18	,7547	,7981	,8402
	Efectividad Lenguaje	2575	,7489	,10076	,29	1,07	,6814	,7453	,8208

**Fuente:** elaboración propia, utilizando la base de datos (2010b) y el repositorio digital del Icfes (2010c)

La tabla 6 ofrece el número, media y suma de los rangos negativos y de los rangos positivos (las notas al pie de la tabla permiten conocer el significado de los rangos positivos y negativos). También ofrece el número de empates (casos que no son incluidos en el análisis)

y el número total de sujetos. Se aprecia que existen mayores ocasiones en las cuales el %efecSaberPRO es mayor que el %efecSaber11, situación que se tiene en consideración a través de la proporción de estudiantes que igualan o superan su desempeño relativo.

**Tabla 6.** Número, media, suma de rangos negativos y rangos positivos, número de empates y número total de sujetos por regiones Corpes para las componentes “Comprensión Lectora” y “Lenguaje”

Ranks					
CORPES			N	Mean Rank	Sum of Ranks
AMAZONIA	Efectividad Lenguaje -	Negative Ranks	923 <sup>a</sup>	636,22	587232,50
	Efectividad	Positive Ranks	240 <sup>b</sup>	373,47	89633,50
	Comprensión Lectora	Ties	1 <sup>c</sup>		
		Total	1164		
CENTRO ORIENTE	Efectividad Lenguaje -	Negative Ranks	54646 <sup>a</sup>	48085,99	2627707279
	Efectividad	Positive Ranks	35496 <sup>b</sup>	40430,69	1435127875
	Comprensión Lectora	Ties	55 <sup>c</sup>		
		Total	90197		
COSTA ATLÁNTICA	Efectividad Lenguaje -	Negative Ranks	18691 <sup>a</sup>	15208,93	284270096,00
	Efectividad	Positive Ranks	9374 <sup>b</sup>	11688,29	109566049,00
	Comprensión Lectora	Ties	15 <sup>c</sup>		
		Total	28080		
OCCIDENTE	Efectividad Lenguaje -	Negative Ranks	33847 <sup>a</sup>	28349,60	959549067,00
	Efectividad	Positive Ranks	18364 <sup>b</sup>	21970,77	403471299,00
	Comprensión Lectora	Ties	18 <sup>c</sup>		
		Total	52229		
ORINOQUÍA	Efectividad Lenguaje -	Negative Ranks	1708 <sup>a</sup>	1411,81	2411374,50
	Efectividad	Positive Ranks	866 <sup>b</sup>	1042,32	902650,50
	Comprensión Lectora	Ties	1 <sup>c</sup>		
		Total	2575		

a. Efectividad Lenguaje < Efectividad Comprensión Lectora

b. Efectividad Lenguaje > Efectividad Comprensión Lectora

c. Efectividad Lenguaje = Efectividad Comprensión Lectora

**Fuente:** elaboración propia, utilizando la base de datos (2010b) y el repositorio digital del Icfes (2010c).

**Tabla 7.** Estadístico de Wilcoxon (Z) y su nivel crítico bilateral (sig. Asintót. Bilateral) por regiones Corpes para las componentes “Comprensión Lectora” y “Lenguaje”.

Test Statistics <sup>b</sup>			
CORPES			Efectividad Lenguaje - Efectividad Comprensión Lectora
AMAZONIA	Z		-21,717 <sup>a</sup>
	Asymp. Sig. (2-tailed)		,000
CENTRO ORIENTE	Z		-76,323 <sup>a</sup>
	Asymp. Sig. (2-tailed)		,000
COSTA ATLÁNTICA	Z		-64,358 <sup>a</sup>
	Asymp. Sig. (2-tailed)		,000
OCCIDENTE	Z		-80,732 <sup>a</sup>
	Asymp. Sig. (2-tailed)		,000
ORINOQUÍA	Z		-20,005 <sup>a</sup>
	Asymp. Sig. (2-tailed)		,000

a. Based on positive ranks.

b. Wilcoxon Signed Ranks Test

**Fuente:** elaboración propia, utilizando la base de datos (2010b) y el repositorio digital del Icfes (2010c)

La tabla 7 muestra el estadístico de Wilcoxon (Z) y su nivel crítico bilateral (sig. Asintót. Bilateral). Puesto que el valor del nivel crítico (0,000 para todas las regiones Corpes) es menor que 0,05, se rechaza la hipótesis de igualdad de promedios y, por tanto, se puede concluir que las variables comparadas difieren significativamente, mostrando una vez más la influencia del proceso formativo.

### La prueba de signos

La tabla 8 permite apreciar el estadístico Z (pues el tamaño muestral es mayor de 25)

y su correspondiente nivel crítico bilateral (sig. Asintót. Bilateral). Debido a que el valor del nivel crítico (0,000) es menor que 0,05, se puede rechazar la hipótesis de igualdad de promedios y, por tanto, concluir que las variables comparadas (rendimiento en “Comprensión Lectora” y rendimiento en “Lenguaje”) difieren. Vale la pena resaltar que este análisis solamente toma en cuenta los signos de las diferencias, más no la magnitud de las mismas y en ese sentido las dos pruebas aplicadas son mucho más robustas cuando existe una medición en términos de los rangos contemplados por las diferencias.

**Tabla 8.** Estadístico Z y su correspondiente nivel crítico bilateral (sig. Asintót. Bilateral) por regiones Corpes para las componentes “Comprensión Lectora” y “Lenguaje”

Test Statistics <sup>a</sup>		
CORPES		Efectividad Lenguaje - Efectividad Comprensión Lectora
AMAZONIA	Z	-19,998
	Asymp. Sig. (2-tailed)	,000
CENTRO ORIENTE	Z	-63,780
	Asymp. Sig. (2-tailed)	,000
COSTA ATLÁNTICA	Z	-55,609
	Asymp. Sig. (2-tailed)	,000
OCCIDENTE	Z	-67,756
	Asymp. Sig. (2-tailed)	,000
ORINOQUÍA	Z	-16,576
	Asymp. Sig. (2-tailed)	,000

a. Sign Test

**Fuente:** elaboración propia, utilizando la base de datos (2010b) y el repositorio digital del Icfes (2010c)

## CONCLUSIONES

El estudio realizado presentó una propuesta metodológica que permitiese estimar la diferencia entre los desempeños relativos de los estudiantes antes y después de un proceso formativo universitario, tomando como

referente el Examen para el Ingreso a la Educación Superior (Saber11) y el Examen de Estado de Calidad de la Educación Superior (SaberPRO); haciendo un análisis comparativo por regiones de planificación Corpes.

Teniendo en cuenta que dichos mejoramientos son de naturaleza multicausal y



este hecho limita el alcance del presente estudio (Saavedra, 2011), en las regiones Corpes analizadas se observó un notable mejoramiento del porcentaje de efectividad máxima antes y después de la formación universitaria (%efecSaberPRO fue superior a %efecSaber11).

De la misma manera, se llegó a la conclusión de que no existe una asociación muy fuerte entre los porcentajes antes y después, situación que podría considerarse adecuada en un proceso de intervención educativa, pues de alguna manera los estudiantes no “cargan con un destino académico” y las instituciones a través de sus programas podrían generar VA en sus desempeños en el contexto que los rodea, incorporando, lógicamente, un beneficio social intrínseco.

Cuando se realizan pruebas estadísticas para determinar si se podría considerar un progreso significativo en los porcentajes antes y después del proceso formativo (en las pruebas empleadas), se llegó a la conclusión de que los porcentajes de efectividad demuestran diferencias significativas y, por tanto, la posición relativa en la prueba SaberPRO es significativamente mejor a la que tenían antes de ingresar a la formación universitaria (prueba Saber11).

Para finalizar, vale la pena resaltar que esta propuesta académica se encuentra en construcción continua, y se buscan estrategias que permitan realizar miradas de los procesos formativos desde una perspectiva integral (y con proyección), lo cual se construye día tras día en una importante fuente de análisis de los procesos de mejoramiento continuo adelantados en las IES. Un análisis siguiente se deberá concentrar en un

estudio más profundo por departamentos y programas (en sus niveles técnicos, tecnológicos y profesionales).

## REFERENCIAS

- Acosta Medina, A. (s.f.). *Regiones Administrativas de Planeación*. Recuperado el 15 de febrero de 2012, de [http://www.sogeo-col.edu.co/documentos/reg\\_activas.pdf](http://www.sogeo-col.edu.co/documentos/reg_activas.pdf)
- Amat, O. (2000). *EVA, Valor Económico Agregado*. Bogotá: Norma.
- Encuesta Internacional de Alfabetización de Adultos [IALS]. (1998). Santiago de Chile, Chile.
- Froemel Andrade, J. E. (2003). *Evaluación de la calidad de la educación con equidad: el modelo de valor agregado*. Recuperado el 19 de julio de 2012, de <http://www.personaysociedad.cl/wp-content/uploads/2011/03/11-Froemel.pdf>.
- Galvis, L. A. y Bonilla, L. (2010). *Desigualdades en la distribución del nivel educativo de los docentes en Colombia*. Recuperado el 26 de agosto de 2012, de <http://www.banrep.gov.co/documentos/publicaciones/regional/documentos/DTSER-151.pdf>
- Galvis, L. A. y Meisel Roca, A. (2010). *Persistencia de las desigualdades regionales en Colombia: Un análisis espacial*. Recuperado el 26 de agosto de 2012, de <http://www.banrep.gov.co/documentos/publicaciones/regional/documentos/DTSER-120.pdf>
- Goldschmidt, P. (18 y 19 de octubre de 2006). De estratificación a crecimiento y modelos de valor agregado. Conectando investigación y práctica. *Seminario Internacional: Medición de Aprendizajes y Valor Agregado en el Sistema Escolar..* Santiago de Chile, Chile.

- Instituto Colombiano para la Evaluación de la Educación [ICFES] (2010a). *SaberPRO-Examen de Estado de la Calidad de la Educación Superior. Qué se evalúa y cómo se presentan los resultados*. Recuperado el 25 de febrero de 2012, de [http://www.icfes.gov.co/index.php?option=com\\_content&task=view&id=568&Itemid=1061](http://www.icfes.gov.co/index.php?option=com_content&task=view&id=568&Itemid=1061).
- Instituto Colombiano para la Evaluación de la Educación [ICFES] (2010b). *Base de datos SB11-SBPRO-CRUCÉ*. Bogotá.
- Instituto Colombiano para la Evaluación de la Educación [ICFES] (2010c). *Repositorio digital del Instituto Colombiano para la Evaluación de la Educación*. Bogotá.
- Jacob, B. A., Kane, T. J., Rockoff, J. E. y Staiger, D. O. (2009). *Can you recognize an effective teacher when you recruit one?* Ann Arbor: Center for Local State and Urban Policy [CLOSUP], University of Michigan.
- Lind, D. A., Marchal, W. G. y Watchen, S. A. (2008). *Estadística aplicada a los negocios y la economía*. Bancroft: The University of Toledo.
- Martínez Sánchez, A. (2001). *Competencia laboral y su contexto*. Recuperado el 11 de agosto de 2012, de [http://www.sistemas-deconocimiento.org/Portal/p\\_csc3.html](http://www.sistemas-deconocimiento.org/Portal/p_csc3.html)
- Meckes, G. L. y Ramírez, M. J. (2006). *Simce: Evaluando el aporte de medir valor agregado*. Santiago de Chile: Ministerio de Educación.
- Medina Durango, C. A. (agosto de 2010). Algunos avances y retos de la educación superior en Colombia. *Encuentro Nacional de la Asociación Colombiana de Universidades [ASCUN]*. Bogotá, Colombia.
- Medina Durango, C. A. y Posso Suárez, C. M. (2010). Technical change and polarization of the labor market: evidence for Brazil, Colombia and Mexico. *Borradores de Economía*, 614, 1-51.
- Ramírez Gomez, M. y Forero Ramírez, N. (2008). Determinantes de los ingresos laborales de los graduados universitarios en Colombia, un análisis a partir de la herramienta de seguimiento a graduados. *Revista de Economía del Rosario*, 11(1), 61-103.
- Ramírez Gómez, M., Vélez, M., Latorre, C., Rey, M., Pombo, C et al. (2005). *Hacia un Sistema Nacional de Información de la Educación Superior*. Bogotá: Universidad del Rosario.
- Ray, A. (18 y 19 de octubre de 2006). Expertos internacionales analizaron los principales avances y limitaciones en el uso de la metodología de valor agregado. *Seminario Internacional: Medición de Aprendizajes y Valor Agregado en el Sistema Escolar*. Santiago de Chile, Chile.
- Rocha de la Torre, A., Olaya Castro, A., Pedraza Daza, P., Cuchimaque, E., Verano Gamboa, L. et al. (2011). *Nuevo Examen de Estado para el Ingreso a la Educación Superior. Cambios para el siglo XXI. Propuesta general*. Recuperado el 27 de febrero de 2012, de [http://www.icfes.gov.co/index.php?option=com\\_docman&task=docview&gid=1200](http://www.icfes.gov.co/index.php?option=com_docman&task=docview&gid=1200)
- Rockoff, J. E. (2004). The impact of the individual teachers on student achievement: evidence from panel data. *The American Economic Review*, 94(2), 247-286.
- Saavedra, J. E. (2011). Resource constraints and educational attainment in developing countries: Colombia 1945-2005. *Journal of Development Economics*, 99(1), 80-91.
- Sammons, P., Thomas, S. y Mortimore, P. (1997). *Forging Links. Effective schools, effective departments*. London: Paul Chapman.
- Sammons, P., Thomas, S., Owen, C. y Pennell, H. (1994). *Asssesing school effectiveness: developing measures to put school*

- performance in context*. London: Office for Standards in Education [OFSTED].
- Shavelson, R. J. (24 de noviembre de 2010). Importancia de las competencias genéricas en la educación media y superior. Conferencia presentada en el Instituto Colombiano para la Evaluación de la Educación [ICFES]. Bogotá, Colombia.
- Shavelson, R. J. (5 y 6 de noviembre de 2008). La medición como una manera de expresar la evaluación del estudiante: sus posibilidades y limitaciones. *Foro Nacional de Evaluación del Aprendizaje en la Educación Superior*. Bogotá, Colombia.
- Thomas, S. y Mortimore, P. (1996). Comparison of value added models for secondary school effectiveness. *Research Papers in Education*, 1(11), 5-33.
- Universidad de Cádiz (s.f.). *Guía para el análisis de datos*. Recuperado de <http://www2.uca.es/serv/ai/formacion>
- Vargas Cuevas, F. E. (2006). *Valor agregado de la educación universitaria* (Tesis de maestría) Santiago de Chile: Universidad de Chile.
- Vásquez Rizo, F. E. (2010). Sistemas de valor basados en conocimiento. En D. Quiroga Parra, F. E. Vásquez Rizo, H. Montaña Motato, D. Espinosa Correa, B. E. Hernández Arias y P. A. Gutiérrez Morales. *La gestión del conocimiento y las tecnologías de la información y la comunicación en las organizaciones* (pp. 87-124). Cali: Universidad Autónoma de Occidente.