

Vol. 21, 2019/e28

Los resultados de la prueba Saber 11 de Ciencias Sociales y las opiniones de los estudiantes: convergencias y divergencias

Results of the Saber 11 Test of Social Sciences and Students Opinions: Convergences and Divergences

Nancy Palacios Mena (1) n.palaciosm@uniandes.edu.co
Miguel Armando Rodríguez Márquez (2) marodriguezm@ut.edu.co

(1) Universidad de Los Andes

(2) Universidad del Tolima

(Recibido: 1 de agosto de 2017; Aceptado para su publicación: 12 de febrero de 2018)

Cómo citar: Palacios, N. y Rodríguez, M. A. (2019). Los resultados de la prueba Saber 11 de Ciencias Sociales y las opiniones de los estudiantes: convergencias y divergencias. *Revista Electrónica de Investigación Educativa*, 21, e28, 1-18.
doi:[10.24320/redie.2019.21.e28.2116](https://doi.org/10.24320/redie.2019.21.e28.2116)

Resumen

El objetivo de este artículo es analizar, por un lado, los resultados de las pruebas de Ciencias Sociales del examen de Estado del año 2016 en Colombia y, por el otro, las opiniones de los estudiantes sobre la preparación que reciben para dicha prueba en el marco de los procesos de evaluación de la asignatura. Se utilizó una metodología cuantitativa con información de dos fuentes: el Instituto Colombiano para la Evaluación de la Educación y una encuesta. Los datos se sistematizaron utilizando el programa estadístico Infostad Profesional. Se determinó que, si bien estas pruebas describen el logro de niveles particulares de desempeño de los estudiantes en el área de Ciencias Sociales, es necesario que estos resultados puedan ayudar a comprender aspectos curriculares y prácticas de enseñanza; esto implica un desplazamiento en la pregunta “cuánto saben los estudiantes” a “qué saben” y “qué son capaces de hacer”, y centrar la reflexión en los usos pedagógicos de los resultados de los exámenes.

Palabras clave: Educación secundaria, evaluación del aprendizaje, desarrollo de competencias, exámenes estandarizados, Ciencias Sociales.

Abstract

The objective of the article is to analyze, on one hand, the results of the Social Sciences tests of the 2016 State Examination in Colombia and on the other, the students' opinions on the preparation they receive for said test within the framework of the evaluation processes of the subject. For this, a quantitative methodology was used with information from two sources: the Colombian Institute for the Evaluation of Education (ICFES) and a survey. The data was systematized using the Infostad Professional statistical program. It was determined that although these tests describe the achievement of particular levels of performance of students in the area of Social Sciences, it is necessary that these results can help to understand curricular aspects and teaching practices; this implies a shift from asking how much the students know, to asking what they know and what they are capable of doing and focusing the reflection on the pedagogical uses of the results of the exams.

Keywords: Secondary education, student evaluation, skills development, standardized tests, Social Sciences.

I. Introducción

Este documento forma parte de estudios que problematizan la relación entre la evaluación como un indicador de la calidad de la educación y las percepciones que tienen los estudiantes sobre dicho proceso. El objetivo del artículo es analizar los resultados de las pruebas de Ciencias Sociales del examen de Estado del año 2016 en Colombia y las opiniones de los estudiantes sobre la preparación que reciben para dicha prueba en el marco de los procesos de evaluación de la asignatura.

El documento es uno de los productos de la investigación sobre la enseñanza de las Ciencias Sociales en la escuela secundaria de hoy, que toma como referente empírico tres instituciones educativas de la ciudad de Ibagué (Colombia). Con el análisis de los resultados de las pruebas Saber 11¹ obtenidos por los estudiantes en el área señalada, sus opiniones sobre la preparación que reciben para dicho examen, y los procesos de evaluación, se espera establecer qué aprendizajes evidencian los resultados de las pruebas de Ciencias Sociales y Competencias Ciudadanas.

El estudio surge de la necesidad de fomentar investigaciones que ilustren los usos pedagógicos que se le pueden dar a los resultados de la prueba Saber, especialmente en aspectos como la pertinencia de los currículos, las estrategias pedagógicas, las actividades de aprendizaje y los procesos de evaluación que se están implementando en las aulas de clases. Este tipo de trabajos también aportan a los debates entre quienes diseñan las políticas educativas y quienes las implementan, con el objetivo de alcanzar consensos sobre estrategias de enseñanza y evaluación que redunden en el mejoramiento del sistema educativo colombiano. (Rojas, 2014).

La principal finalidad de las pruebas de Estado es medir la madurez académica del estudiante, los conocimientos y capacidades adquiridos en su formación durante la primaria y el bachillerato; asimismo, buscan evaluar su capacidad para seguir con éxito la formación universitaria. El impacto y las implicaciones de los resultados de estas pruebas han ido en aumento en los últimos años, pues además de “clasificar a los alumnos, clasifica a las instituciones en las categorías A, B, C”, determinan qué estudiantes pueden acceder al programa de becas de educación superior del Ministerio de Educación Nacional (MEN), denominado “Ser Pilo Paga”, y es uno de los criterios de calidad que hace acreedores a incentivos económicos a maestros y directivos de las instituciones educativas públicas. En la prueba de Ciencias Sociales y Competencias Ciudadanas se examina si los estudiantes son capaces de buscar información, de operar con la que poseen, y se evalúan los niveles de comprensión que alcanzan respecto a procesos sociales pasados y actuales, la tabla I muestra la estructura de dichas pruebas.

La actual estructura de las pruebas (ver tabla I) evidencia que en la determinación de las competencias incluidas, su conceptualización y los criterios de evaluación, hay un claro propósito de articular dichas pruebas con los lineamientos curriculares de los Estándares Básicos de Competencias en Ciencias Sociales y Ciudadanas.² Asimismo, ilustra que en la definición de las competencias son fundamentales la comprensión de fenómenos sociales, la apropiación de conceptos particulares de las Ciencias Sociales, la aprehensión de actitudes y valores que hagan explícita la adquisición de cierto nivel de pensamiento crítico frente a la realidad cotidiana, y el desarrollo de habilidades y destrezas propias de la labor de un científico social (MEN-ICFES, 2013).

¹ Los resultados de las pruebas Saber 11 en la educación media, las pruebas Saber de 3, 5 y 9 en la educación básica y las pruebas Saber Pro para la educación superior son indicadores para medir la calidad de la educación en Colombia. Las pruebas son producidas y calificadas por el Instituto Colombiano para la Evaluación de la Educación (ICFES), entidad que forma parte del Ministerio de Educación Nacional (MEN).

² Los estándares básicos de competencias constituyen uno de los parámetros de lo que todo niño y joven debe saber y saber hacer para lograr el nivel de calidad esperado a su paso por el sistema educativo; la evaluación externa e interna es el instrumento por excelencia para saber qué tan lejos o tan cerca se está de alcanzar la calidad establecida con los estándares (MEN, 2006, p. 9).

Tabla I. Estructura de las pruebas de Ciencias Sociales y Competencias Ciudadanas

Competencia	Estándares básicos de Ciencias Sociales y de Competencias Ciudadanas
Pensamiento social	- Relaciones espaciales y ambientales Convivencia y paz
Interpretación y análisis de perspectivas	- Relaciones ético-políticas
Pensamiento reflexivo y sistémico	- Participación y responsabilidad democrática
	- Relaciones con la historia y las culturas
	- Pluralidad, identidad y valoración de las diferencias

Fuente: MEN-ICFES (2013).

Las pruebas de Ciencias Sociales otorgan una calificación numérica y las preguntas que involucran algunos de los ámbitos de Formación para la Ciudadanía dan otra calificación que se presenta con una valoración numérica aparte de la señalada para Ciencias Sociales y para Formación para la Ciudadanía (MEN-ICFES, 2015). La descripción de los niveles de desempeño es fundamental porque, como se indicaba, muestran qué saben y qué no saben hacer los estudiantes y en qué tienen fortalezas y deficiencias.

Como se evidencia en el contenido de la tabla II, las pruebas contemplan un sistema de clasificación de los estudiantes en cuatro niveles, siendo 1 el más bajo y 4 el más alto. Cada nivel tiene unos criterios de clasificación que indican lo que saben y saben hacer los estudiantes. La clasificación del alumnado por niveles de desempeño ha adquirido una gran relevancia por cuanto el reto de los establecimientos es sacar a todos sus estudiantes del rango 1 y tratar de llevarlos del rango 2 al 3 y de éste al 4.

Tabla II. Niveles de desempeño y criterios de clasificación de la prueba de Ciencias Sociales

Nivel	Rango	Descripción
1	0-40	El estudiante que se ubica en este nivel podría reconocer algunos derechos ciudadanos en situaciones sencillas. Podría reconocer factores que generan un conflicto y creencias que explican algunos comportamientos.
2	41-55	El estudiante que se ubica en este nivel reconoce deberes del Estado colombiano y situaciones de protección o vulneración de derechos en el marco del Estado social de derecho.
3	56-70	Además de lo descrito en el nivel 2, el estudiante que se ubica en este nivel identifica prejuicios o intenciones contenidos en una afirmación y reconoce las dimensiones e intereses involucrados en un problema o alternativa de solución.
4	71-100	Además de lo descrito en los niveles 2 y 3, el estudiante que se ubica en este nivel conoce algunas disposiciones de la Constitución Política de Colombia que posibilitan la participación ciudadana y el control a los poderes públicos.

Fuente: MEN-ICFES (2016).

El desarrollo de los estudios sobre evaluación, y más concretamente sobre los resultados de pruebas estandarizadas nacionales e internacionales, hace parte de las líneas de acción que en las últimas dos décadas han orientado las reformas educativas en el ámbito mundial: la ampliación de la cobertura educativa, el acceso universal y gratuito a la educación primaria y secundaria, y el logro de mayores aprendizajes por parte de los estudiantes. Según Ferrer y Arregui (2006) dichas líneas de acción surgen en el contexto de un renovado interés por determinar la calidad de la educación que se imparte en los diferentes países, en un momento político y social donde el debate sobre cuál debe ser el centro del proceso educativo tiene diferentes formas de interpretación en los ámbitos académicos y en la opinión pública.

Estamos ante la búsqueda de relaciones más estrechas entre la educación, la productividad y el mercado; en ese escenario se hace prioritario que los niños y jóvenes desarrollen competencias y destrezas producto de los aprendizajes que han adquirido en la escuela, y que deberían ser evidenciables en el trabajo y en la vida diaria.

Dentro de este marco de interpretación se ha multiplicado el uso de indicadores y estándares en gran parte de los países de América Latina, los cuales, además de diagnosticar los niveles de aprendizaje del alumnado, permiten determinar el nivel de responsabilidad de los alumnos, los profesores y los establecimientos educativos respecto a dichos niveles. Además de ello, estos diagnósticos han estado acompañados de incentivos económicos y reconocimientos honoríficos. Una de las conclusiones a las que llegan Ferrer y Arregui (2006) es que la realización de evaluaciones comparativas y cuantitativas de los logros de aprendizaje de los alumnos ha llevado a que los países que no disponían de sistemas de evaluación los desarrollaran junto con el diseño e implementación de las reformas educativas de los años ochenta y noventa.

En misma línea argumentativa, Ravela, Wolfe, Valverde y Esquivel (2006) afirman que los esfuerzos por institucionalizar sistemas de evaluación en distintos países de América Latina después de 1992, obedecen a la opinión compartida que planteaba que los resultados del proceso educativo no eran claros, puesto que lo que las familias y la sociedad sabían sobre lo que se enseñaba en los centros educativos provenía de fuentes como las tareas que les asignaban los maestros para realizar en la casa o de los informes académicos y de comportamiento entregados a los padres de familia; sin embargo, se carecía de una visión clara sobre lo que los niños y jóvenes estaban aprendiendo.

Según estos autores, en el caso de los maestros, si bien se tenía una visión general respecto a lo que los alumnos eran capaces de hacer, no se tenía una referencia externa con base en indicadores claros y observables acerca de los conocimientos y competencias que lograban adquirir los niños. En el caso de las autoridades educativas, la necesidad era la de unos datos que proporcionaran información más allá de las cifras sobre matrícula, cobertura, repetición y deserción. Empero, como se señaló anteriormente, se requerían datos que dieran información de importancia estratégica para la gestión educativa, como por ejemplo, respecto a los conocimientos y a las competencias para el desempeño social de los estudiantes.

Así, según Ravela et al. (2006), las necesidades de las familias, de los maestros, de los encargados de las políticas públicas en educación y de la sociedad en general, justificaron la creación e institucionalización de sistemas nacionales de evaluación de la educación, los cuales son definidos por Perassi (2008) como organizaciones y articulaciones relativamente complejas, orientadas a evaluar los aprendizajes a nivel nacional con dos funciones fundamentales, una técnica y otra política. La primera, relacionada con la elaboración de instrumentos, validación de criterios, aplicación de pruebas, procesamiento de datos, interpretación de los resultados, comparabilidad de estos y elaboración de los informes de resultados. La segunda, relacionada con la presentación y la comunicación de los resultados, uso de estos, el impacto que dichos resultados provocan en los distintos actores, la relación existente entre la información que surge de los sistemas nacionales de evaluación y la toma de decisiones.

Trabajos como los de Locatelli y Couto Andrade (2001), Ferrer (2005) y Tiana Ferrer (2003) coinciden en que todos los sistemas de educación tuvieron como principal objetivo determinar el estado de la situación del sistema educativo, bajo la premisa fundamental de que la información obtenida en la evaluación permitiría orientar acciones para mejorar la calidad de la educación en cada uno de los países. Hay que decir, sin embargo, que desde algunos sectores de la academia y desde sectores sindicales de profesores se han hecho fuertes críticas a las acciones desarrolladas desde los sistemas nacionales de evaluación; para estos sectores, lejos de contribuir al mejoramiento de la calidad de la educación, las acciones de estos organismos se han orientado a consolidar políticas neoliberales orquestadas desde organizaciones multilaterales que han afectado las condiciones de trabajo del magisterio y que han llevado a la adopción de modelos que afianzan el capitalismo, como los modelos de educación por competencias.

Desde la perspectiva de Perassi (2008):

La acción de los sistemas de evaluación, sobre los resultados de los exámenes estandarizados nacionales, se ha centrado en exhibir y premiar a las escuelas cuyos alumnos obtienen buenos puntajes, presentándolas como las instituciones escolares de mejor calidad. (p. 20).

Para Perassi (2008), la utilización de los resultados de las evaluaciones de los estudiantes como indicador motivo para el otorgamiento de reconocimientos públicos y de incentivos económicos, tiene por lo menos dos graves consecuencias:

Por un lado, desvirtúa el sentido esencial de la evaluación instalando una fuerte presión desde el poder central para que las escuelas alcancen los primeros puestos (...) y por otro lado, reduce la complejidad del concepto de calidad educativa a un puntaje que refleja sólo parcialmente una dimensión de la misma. (p. 20).

Sobre la presentación de los resultados de las pruebas nacionales (motivo de este artículo), Montoya, Perusia y Vera (2001) encontraron que los resultados más socializados son los de las áreas de matemáticas y lenguaje. De igual manera, se evidenció que mucha de la información se queda en los directivos docentes y no es conocida ni usada con fines pedagógicos y didácticos por parte de los docentes; de allí que estos estudios recomiendan generar espacios de discusión entre los productores de información, los supervisores de centros educativos y los planificadores y capacitadores de docentes; sin embargo, trabajos como los de Cueto (2005) subrayan que entre los usos de las evaluaciones nacionales se pueden descartar varias modalidades:

La utilización de las pruebas para la selección de escuelas que dado su bajo rendimiento requieren mayor atención (por ejemplo, el programa 900 escuelas en Chile). Las pruebas han servido en algunos países para evaluar el impacto de determinadas políticas y programas. Adicionalmente ha generado la necesidad de clarificar los requerimientos curriculares; este movimiento estaría orientándose paulatinamente hacia la definición de estándares de rendimiento escolar (hay algunas experiencias en Centroamérica al respecto). Por último, se debe descartar que varios de los países de la región hayan usado las pruebas para diseñar materiales y capacitaciones orientadas a docentes (p. 13).

Desde la postura de Cueto (2005), son necesarios trabajos en los que se haga difusión de estudios empíricos sobre los resultados de las pruebas nacionales, por cuanto pueden constituirse en herramientas que permiten que dichos resultados sean entendidos y discutidos. En esta misma línea de argumentación, Perassi (2008) sostiene que hay necesidad de continuar y multiplicar las investigaciones en torno a los sistemas nacionales de evaluación, sus políticas, sus acciones, el papel que están teniendo en los sistemas educativos las pruebas nacionales y las internacionales, el impacto de sus resultados, así como la ampliación de la reflexión que se produce acerca de los factores asociados al aprendizaje.

Teniendo en cuenta que en Colombia, como en otros países, los resultados de los exámenes estandarizados cobran cada día más importancia como indicadores para medir la calidad de la educación, para clasificar a las instituciones educativas de educación primaria y secundaria, para permitir o no el ingreso a las instituciones de educación superior y para el otorgamiento de estímulos económicos y reconocimientos honoríficos a estudiantes, profesores, y establecimientos que obtienen altos promedios; el artículo compara dichos resultados en tres secundarias colombianas y las opiniones de los estudiantes sobre la preparación que reciben para obtener buenas calificaciones en la prueba de Ciencias Sociales y la coherencia entre lo que se enseña y se evalúa en el área.

II. Método

El objetivo del artículo es analizar los resultados de las pruebas de Ciencias Sociales del examen de Estado del año 2016 en Colombia y las opiniones de los estudiantes sobre la preparación que reciben para dicha prueba en el marco de los procesos de evaluación de la asignatura.

Participaron 264 estudiantes de último grado de tres instituciones secundarias públicas de la ciudad de Ibagué, Tolima, sus edades oscilan entre los 15 y los 18 años (tabla III):

Tabla III. Instituciones educativas participantes en la investigación

Colegio	Año 2015*		Año 2016**	
	Puesto en la Ciudad	Promedio Ciencias Sociales	Puesto en la Ciudad	Promedio Ciencias Sociales
Institución I	25	50	35	52
Institución II	11	52	19	53
Institución III	5	58	3	59

*En 2015 se evaluaron 36 colegios públicos.

**En 2016 se evaluaron 42 colegios públicos.

Fuente: ICES (2016).

Se decidió trabajar con estudiantes de undécimo grado porque son los que presentan las pruebas Saber 11 y porque al encontrarse en el último año de secundaria tienen una visión más completa de lo que ha sido su formación en el área de Ciencias Sociales durante los cinco años de formación primaria y los seis años de bachillerato. Como se ilustra en la tabla III, la selección de las instituciones educativas se realizó de forma aleatoria en la base de datos del ICES, teniendo en cuenta dos criterios: 1) que fueran colegios oficiales y 2) que tuvieran resultados heterogéneos en los resultados de la prueba Saber 11. Del listado se escogió un colegio ubicado en uno de los primeros puestos, uno ubicado hacia la mitad del listado y uno más de la parte inferior de la tabla, según los promedios obtenidos en las pruebas Saber 11 de 2015 y 2016.

Posterior a la búsqueda aleatoria de las instituciones se les envió una invitación para participar en la investigación, allí se indicaron los objetivos del trabajo y la necesidad de autorización para que los estudiantes participaran en la aplicación de la encuesta y las entrevistas. Asimismo, se diligenciaron los consentimientos informados de los estudiantes y sus padres y se procedió con la aplicación de los instrumentos. Se firmó un compromiso de devolución de los resultados de la investigación mediante la entrega de un documento en el que se presentaran y analizaran los desempeños de cada Institución en la prueba Saber de Ciencias Sociales, dicho documento fue entregado, según lo convenido.

En el artículo se integraron dos tipos de variables: en primer lugar, un grupo de variables socioeconómicas, como el estrato, edad y sexo, con las cuales se buscó tener más conocimiento sobre las condiciones de vida de los estudiantes. En segundo lugar, un grupo de variables que dan cuenta de los resultados de los estudiantes en la prueba estandarizada que aplica el MEN, además de la opinión que tiene el alumnado sobre la articulación de dicha prueba con el proceso de evaluación en la clase de Ciencias Sociales. Las variables de este segundo grupo fueron: los contenidos y las habilidades evaluadas, la coherencia entre lo que se enseña y se evalúa, la articulaciones entre la prueba Saber 11 y las técnicas de evaluación que utilizan los profesores, la medición de la comprensión o de la memorización en las evaluaciones, la claridad y coherencia de los mecanismos de evaluación, las similitudes y diferencias entre los exámenes que hacen los docentes y la prueba Saber y la preparación que los estudiantes reciben en la clase de Ciencias Sociales para dicha prueba.

La información analizada proviene de dos fuentes: 1) las bases de datos del ICES disponibles en la página web de la institución y 2) una encuesta aplicada a la totalidad de los estudiantes de grado undécimo de las tres instituciones participantes (sólo faltaron los estudiantes que el día de la aplicación no fueron al colegio). De la primera se obtuvo el reporte de resultados de la aplicación de las pruebas Saber 11 del 2016 por establecimientos educativos;³ dicho reporte contiene el promedio general de los establecimientos educativos en la prueba, los puntajes obtenidos en cada una de las asignaturas evaluadas, y el porcentaje de estudiantes en cada uno de los niveles de desempeño de dichas asignaturas.

³ En la página web del ICES se encuentra a disposición del público una parte de los resultados de cada una de las áreas evaluadas. El reporte contiene los estudiantes matriculados en cada institución y los que presentaron la prueba, el promedio general, los promedios de las áreas, las desviaciones estándar y los resultados por niveles de desempeño. Los datos para cada institución vienen acompañados por los del país y la entidad certificada de instituciones urbanas, rurales, públicas y privadas, de modo que cada establecimiento pueda compararse con otros.

La encuesta recogió las opiniones de los 264 estudiantes sobre la preparación que les hace el colegio para las pruebas Saber 11 y sobre la evaluación de la clase de Ciencias Sociales. En ella los estudiantes respondieron a preguntas que indagaban aspectos como la verificación de la comprensión de lo enseñado a través de los exámenes, la promoción de la memorización, la similitud de los exámenes de la clase con la prueba Saber 11, la claridad del método de enseñanza desde el inicio del curso y el objetivo de los exámenes de la clase en relación con el mejoramiento de los resultados de la prueba Saber 11, entre otros.

Tabla IV. La población encuestada

Instituciones Educativas	Total estudiantes	Encuestas aplicadas
Institución I	71	66
Institución II	102	95
Institución III	112	103

Fuente: Elaboración propia.

Para el diseño del cuestionario fueron consultados el Modelo de Recolección de Datos de Educación del Departamento Administrativo Nacional de Estadística (DANE) de 2017; la Encuesta de Convivencia Escolar y Circunstancias que la Afectan para Estudiantes de 5o. a 11 de Bogotá (DANE, 2011); el cuestionario de Vallejo y Escudero (1999) para evaluar las actitudes de los estudiantes de la escuela secundaria obligatoria (ESO) hacia las Matemáticas, y el trabajo "Usemos las encuestas a los alumnos para mejorar nuestra docencia" de Barrado, Gallego y Valero-García (1999).

Adicional a lo anterior, se llevó a cabo la aplicación de una prueba piloto con cien estudiantes de los tres colegios participantes en la investigación; este ejercicio permitió ajustar y reformar el cuestionario inicial. Entre los ajustes más significativos se pueden destacar: la reducción de preguntas abiertas, el cambio en las opciones de respuestas a algunas de las preguntas porque generaban ambigüedad a la hora de contestarlas, el reemplazo de algunos términos que los estudiantes no entendían con facilidad; la reducción del tamaño general del cuestionario; la reformulación de algunas preguntas que provocaron confusión en los estudiantes; y la inclusión de preguntas que los estudiantes consideraron importantes. En conclusión, el ejercicio piloto fue muy útil para consolidar y corregir el cuestionario que finalmente se aplicó.

La información de los resultados de las pruebas Saber 11 de cada una de las tres Instituciones fue recopilada, verificada y posteriormente sistematizada, a través de procesamientos estadísticos en diagramas, barras y tablas de contingencia. La sistematización de la información de la encuesta requirió de tres procedimientos. Para empezar, se digitaron las respuestas de los estudiantes; una vez obtenida la información, se construyó una base de datos, y posteriormente, esta fue sometida a un análisis de correspondencias múltiples por conglomerados con el paquete estadístico Infostad Profesional 2017. Como resultado se obtuvo la información de frecuencias para cada variable y los diagramas de dispersión (figuras 5 y 6). Los diagramas permitieron representar gráficamente las variables que se cruzaron y la información de frecuencias proporcionó el número y porcentaje de respuestas dadas por los estudiantes para cada una de las variables.

III. Resultados

3.1 Prueba Saber de Ciencias Sociales

El ejercicio de interpretación de los resultados de las pruebas Saber 11 de 2016 en las instituciones participantes en el estudio incluyó el análisis de los promedios generales (de todas las asignaturas), una lectura comparativa de los promedios de cada asignatura y del porcentaje de estudiantes de cada institución en los cuatro niveles de desempeño, el análisis de las opiniones sobre la preparación de las pruebas Saber 11 y los procesos de evaluación en la clase de Ciencias Sociales. Desde la perspectiva de Ravela (2006a), conocer las mediciones y comprender qué significan los diferentes tipos de datos que ofrecen los informes nacionales es relevante porque de ellos depende la comprensibilidad de la

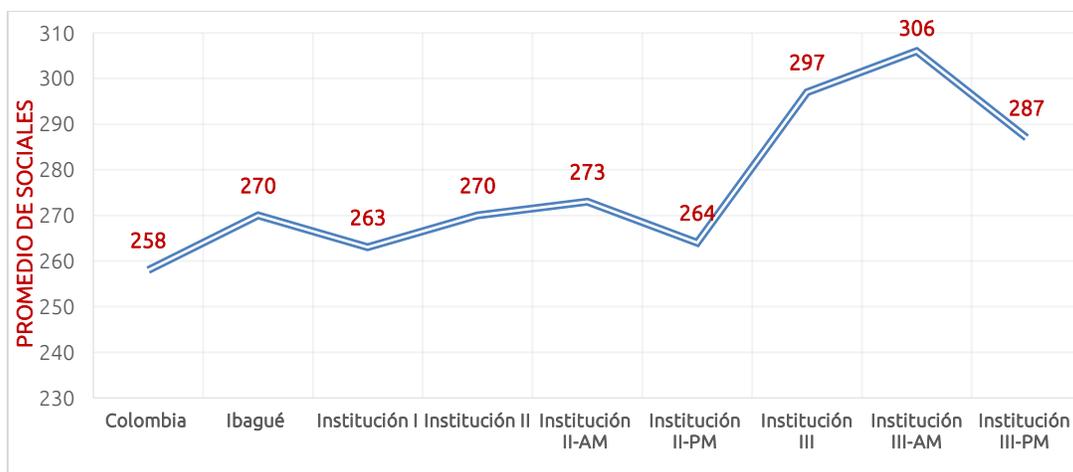
información para los destinatarios, además, está en juego la utilidad de dicha información en función de las necesidades de los usuarios.

3.2 El puntaje global

El puntaje global de un establecimiento es el resultado del promedio ponderado que obtienen los estudiantes en las áreas que son evaluadas en la prueba Saber 11. La ponderación o peso asignado a las pruebas de Matemáticas, Sociales y Ciudadanas, Lectura Crítica y Ciencias Naturales es de tres, mientras que el asignado a la de Inglés es de uno (MEN-ICFES, 2016).

En el reporte de resultados para cada institución, además de su puntaje global se incluye el del país, la ciudad, los colegios privados y los colegios públicos, con el objetivo de que cada colegio se compare con otros. El dato del promedio general es uno de los indicadores que define la clasificación de los planteles.

Según los datos de la figura 1, el puntaje global de Ibagué incluidas las instituciones educativas participantes en el estudio, está por encima del promedio nacional. La institución I lo supera en 5 puntos, la II en 12 y la III en 39. En cuanto a las diferencias entre los tres colegios, la figura 1 ilustra una diferencia de 7 puntos entre la Institución I y la II, pero una diferencia de 34 entre éstas y la Institución III. Nótese que la institución de mejores resultados es la III, y que, aunque la I y la II aparecen clasificadas en B hay entre las dos una diferencia de 7 puntos, evidenciando que los rangos dentro de cuales se ubican los colegios pueden ser amplios, y se pueden encontrar dentro de ellos establecimientos con diferencias significativas.

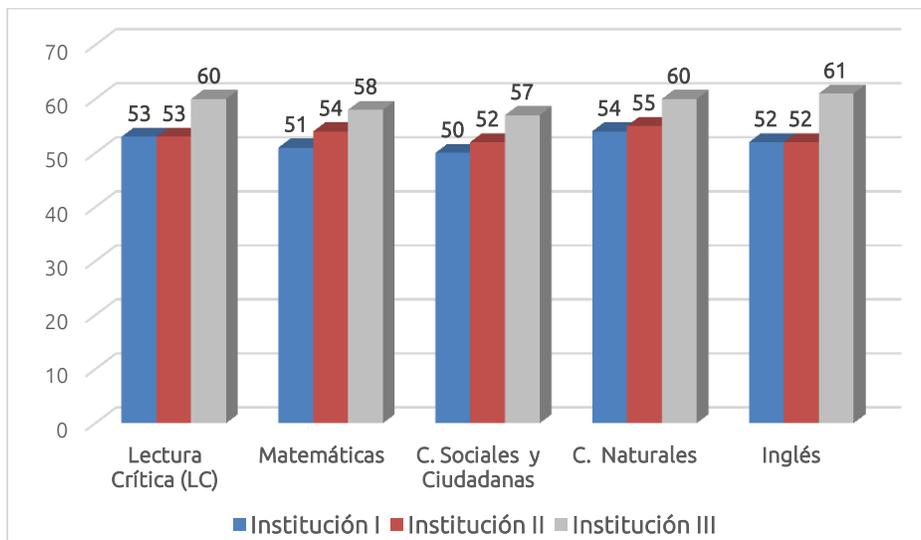


Fuente: MEN-ICFES (2016).

Figura 1. Puntaje global de las pruebas Saber 11, 2016

3.3 Promedio por asignatura

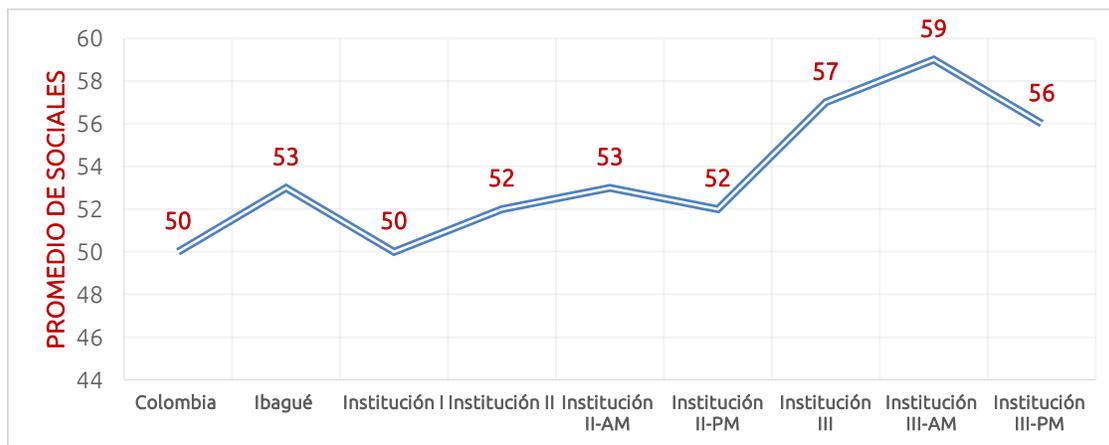
Para nuestro estudio, al analizar la prueba de Ciencias Sociales resultaba útil comparar los resultados de la asignatura con las demás. La figura 2 permite ver dicha comparación. Según el MEN-ICFES (2016), el puntaje promedio expresa el puntaje más representativo de los estudiantes en determinada prueba y se reporta en una escala de 0 a 100 puntos, sin decimales.



Fuente: MEN-ICFES (2016).

Figura 2. Puntaje promedio por asignaturas

La figura 2 muestra que el promedio más bajo de la prueba en las tres instituciones es el del componente de Ciencias Sociales (50 puntos), y en el ámbito nacional este componente y el de Matemáticas son los de más bajo promedio (50). En las tres instituciones los resultados más bajos son los de Ciencias Sociales y competencias Ciudadanas.



Fuente: MEN-ICFES (2016).

Figura 3. Promedio del área de Ciencias Sociales y Competencias Ciudadanas, pruebas Saber 11, 2016

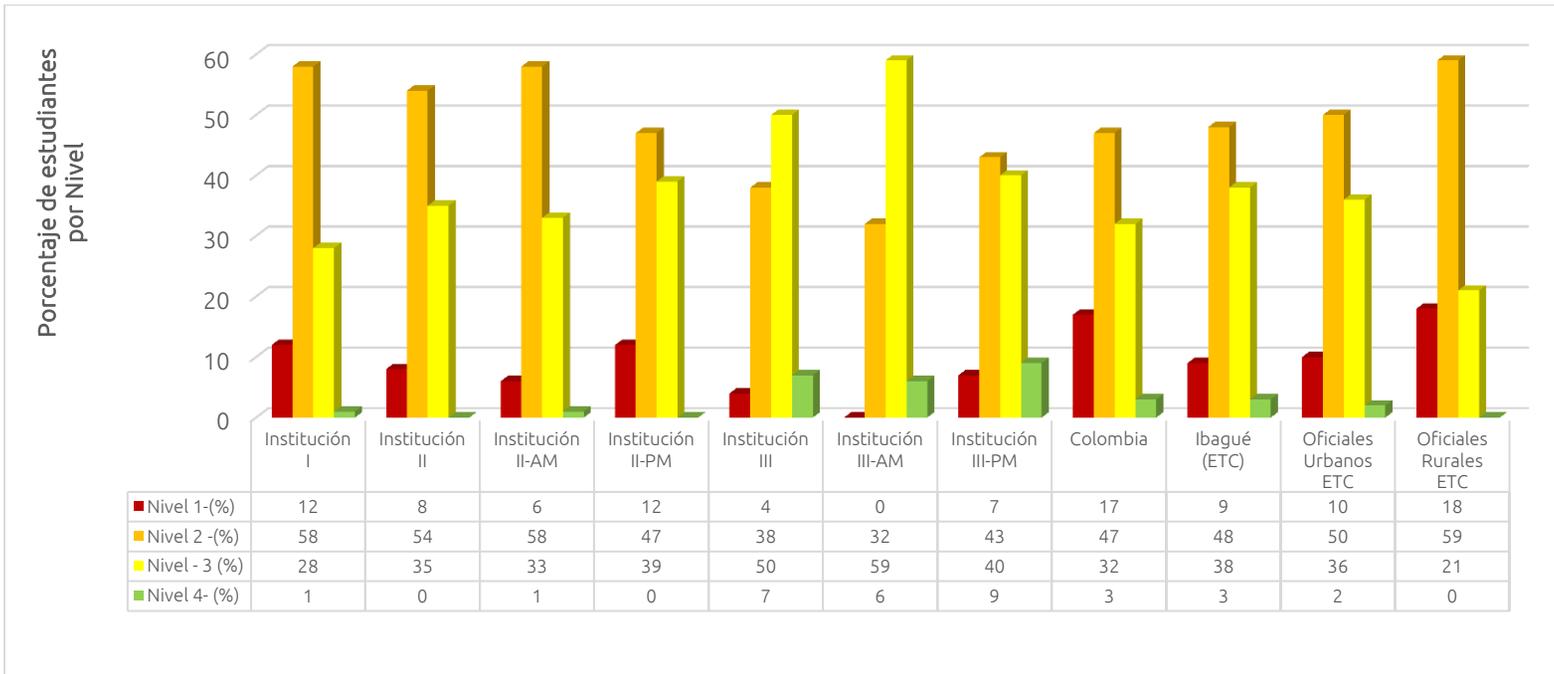
Como se evidencia en la figura 3, no hay grandes diferencias entre el promedio Nacional de Ciencias Sociales y Competencias Ciudadanas y el de las instituciones de Ibagué y las participantes en el estudio. El de la Institución I y el Nacional son de 50 puntos, el de Ibagué 53, la Institución II obtuvo 52 y la III 57, lo que indica que hay sólo una diferencia de 7 puntos entre el menor y el mayor promedio.

En cuanto a la Institución II, su resultado de la prueba de Ciencias Sociales y Competencias Ciudadanas es superior al de Colombia e inferior al general de Ibagué. Por jornadas la tendencia a la homogeneidad se mantiene: hay 1 punto de diferencia que no resulta significativo en el promedio general del área. En el caso de la Institución III, el promedio de la prueba supera el del país y el de los otros establecimientos educativos de Ibagué, tanto públicos como privados. Hay una pequeña diferencia de 3 puntos entre la jornada de la mañana y la jornada de la tarde; sin embargo, esa diferencia no es tan grande como la del promedio general. Este hecho indica que en esta área los resultados no son considerablemente

heterogéneos y que debe ser en otras áreas donde se presentan diferencias más grandes.

3.4 Niveles de desempeño

Los niveles de desempeño describen lo que saben y saben hacer los estudiantes en una determinada prueba de examen. De acuerdo con las especificaciones de los estándares básicos de competencias establecidos por el MEN, para la prueba de Ciencias Sociales se definieron cuatro niveles de desempeño: 1, 2, 3 y 4 (MEN-ICFES, 2006). Cada nivel está definido por un rango en el puntaje, tiene una descripción de los saberes que deben tener los estudiantes y unos criterios de clasificación, por ello, la suma del porcentaje de alumnos en cada nivel no necesariamente da el 100 % (ver tabla II).



Fuente: MEN-ICFES (2016).

Figura 4. Porcentaje de estudiantes por niveles de desempeño en la prueba de Ciencias Sociales y Competencias Ciudadanas, Prueba Saber 11, año 2016

Una lectura global de los resultados por nivel indica que en la prueba de Ciencias Sociales y Competencias Ciudadanas la mayor parte de los estudiantes del país se encuentran en el nivel 2 (ver figura 4). De las tres instituciones del estudio, la III es la que mayor porcentaje de estudiantes tiene en el nivel 3 (50%); la I y la II tienen en este nivel 28% y 35%, respectivamente. En el nivel de menor desempeño la institución que más estudiantes tiene es la I (12%) y la que menos tiene es la III (4%), en el punto intermedio está la Institución II con el 8%. La tendencia general muestra un porcentaje muy mínimo de estudiantes en el nivel más alto de desempeño, el 4: solo el 3% tanto en el país como en Ibagué; la Institución I tiene el 1%, la III el 7% y la II no tiene. Como se había indicado, en la presentación de resultados por institución aparece un estudiante en la jornada de la mañana, pero al unir las dos jornadas deja de ser significativo.

Por niveles de desempeño la Institución I tiene la mayor parte, 48 estudiantes (58%) en el segundo nivel; 23 (28%) en el tercero; 10 (12%) en el primero y solo un estudiante (1%) se ubicó en el 4, el nivel deseado según el MEN. En la Institución II, el 54% (121 estudiantes) se ubicó en el nivel 2, el 35% (78) en el tres, en el nivel más bajo el 8% (18) y en el 4 no hay ninguno. Por jornadas, la tarde concentra 39% (47 estudiantes) entre el segundo y el tercer nivel, 12% en el nivel 1 y ninguno en el nivel 4. La jornada de la mañana tiene 58% en el nivel 2, 33% en el 3, 6% en el 1 y 1% en el nivel 4. La razón por la que el 1% de la jornada de la mañana en el nivel 4 desaparece en el porcentaje general de la institución educativa se debe a que al sumar las dos jornadas el valor no es significativo; es decir, que como presentaron la prueba 264 estudiantes y sólo uno se ubicó en el nivel 4 el valor se considera irrelevante en el resultado general.

En cuanto a la Institución III, en el nivel de menor desempeño el 1 tiene 10 estudiantes (4%), en el nivel 2 hay 92 estudiantes (38%), en el 3 está la mayor parte de los alumnos (121) (79%) y en el nivel 4 se ubicaron 17 (7%). La distribución de los estudiantes por niveles de desempeño en la Institución muestra algunos datos que vale la pena resaltar. La jornada de la mañana concentra la mayor parte entre el nivel 2 y el 3; 32% y 59%, respectivamente; además, no tiene estudiantes en el nivel 1 y 6% se ubica en el nivel 4. La jornada de la tarde es un poco más heterogénea: si bien tiene 9% de sus estudiantes en el nivel 4, a diferencia de la mañana tiene el 7% de los estudiantes en el nivel 1, 43% en el 2 y el 40% en el 3.

3.5 Opinión de los estudiantes sobre la preparación que reciben para presentar la Prueba Saber de Ciencias Sociales

Con el objetivo de contrastar los resultados ya presentados de la prueba Saber 11 y las opiniones de los estudiantes sobre la preparación que reciben para dicha prueba, se hizo una lectura comprensiva de las respuestas dadas por el alumnado a preguntas de la encuesta relacionadas con la evaluación del aprendizaje en la clase de Ciencias Sociales (ver figuras 5 y 6).

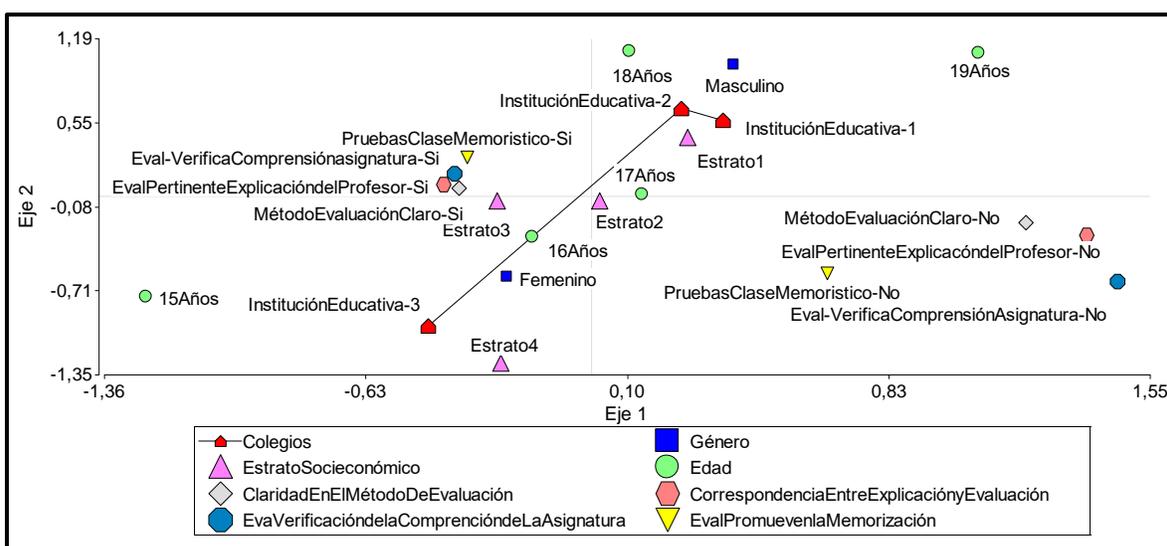


Figura 5. Evaluación de la clase de Ciencias Sociales

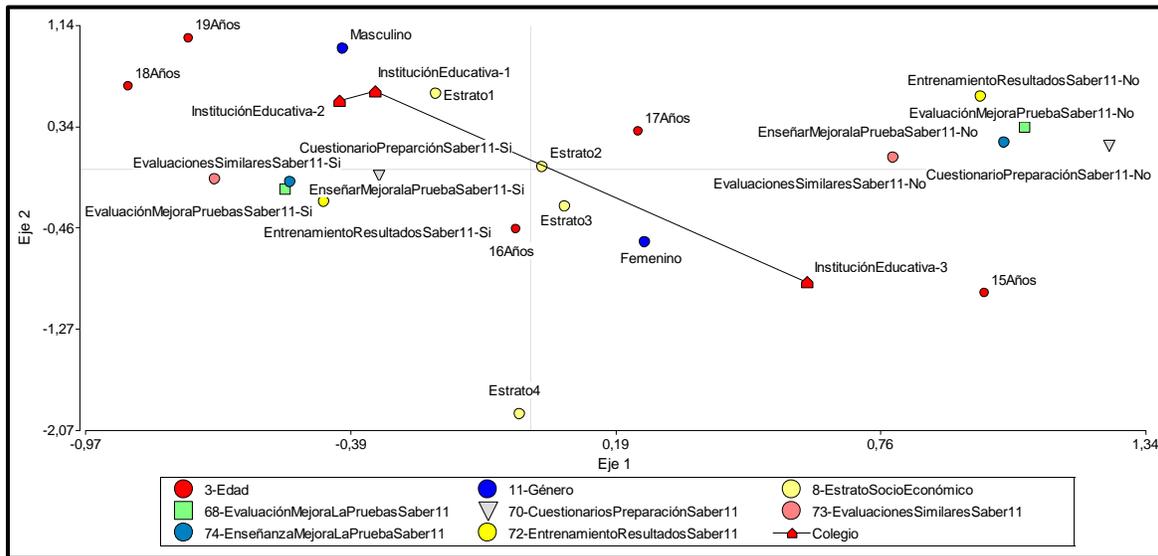


Figura 6. Preparación para las pruebas Saber 11

El eje 1 concentra las respuestas de la opción "No" y el eje 2 las respuestas con la opción "Sí". Los datos de la figura 6 indican que hay más coincidencias entre los estudiantes de la Institución I y II, que a su vez se muestran distantes de la Institución III. Asimismo, hay más coincidencias entre los estudiantes de 16 y 17 años, de los estratos 1, 2 y 3, quienes además tienen notorias diferencias con los del estrato 4. Los resultados indican también que 175 estudiantes (67%) consideran que los resultados de los exámenes promueven la memorización de contenidos; el 79% cree que hay correspondencia entre las explicaciones del profesor y lo que evalúa; 203 (77%) opinaron que los métodos de evaluación son claramente explicados y, por último, 203 estudiantes (77%) consideran que la evaluación verifica la comprensión de la asignatura. Si bien a simple vista pareciera contradictorio que 67% opine que los exámenes promueven la memorización de contenidos y 77% considere que ésta verifica la comprensión de la asignatura, una lectura a las explicaciones dadas al "Sí" y al "No" permite establecer que lo anterior se debe a la forma como los estudiantes conciben qué es comprender. Así, por ejemplo, a la pregunta: "¿Los exámenes de la asignatura verifican la comprensión de lo enseñado? Sí ___ No ___ ¿Por qué?", algunas de las explicaciones dadas son:

- Porque los exámenes son sobre los temas vistos.
- Porque las preguntas no son difíciles.
- Porque muchos ganan los exámenes.
- Porque en los exámenes uno escribe lo que sabe.
- Porque el profesor pregunta lo que ha enseñado.
- Porque los exámenes son sobre las preguntas de los talleres y sobre los mapas que hacemos.
- Porque casi todo mundo pasa y nadie pierde.
- Porque uno en el examen escribe lo que uno piensa.
- Porque en los exámenes hay preguntas de las noticias.
- Porque las preguntas van acorde a los temas.

Estas respuestas evidencian que los estudiantes relacionan la comprensión con lo conocido, con lo sencillo, con lo fácil; para ellos comprender es expresar libremente la opinión sobre un tema en particular.

Llama la atención que siendo los resultados de Ciencias Sociales los más bajos, como se muestra en la figura 2, las opiniones de los estudiantes sean mayoritariamente favorables a la evaluación en el salón de

clases. Estas opiniones parecen corresponder a la relación que establecen entre lo positivo, lo fácil y lo poco complejo que les resulta la asignatura, aunque no haya ni un aprendizaje real de los contenidos ni el desarrollo de competencias, como lo evidencian los resultados del examen.

En el eje 1 se concentran las respuestas para la opción “No” y en el eje 2 las respuestas para la opción “Sí”. Al igual que en la figura 5, en la 6 hay más coincidencias entre las opiniones de los estudiantes de las Instituciones I y II, que entre estos y los de la Institución III. También hay más coincidencias entre los estudiantes de los estratos 2 y 3, con una leve diferencia entre estos y los del estrato 1, y con una marcada diferencia entre todos los anteriores y los del estrato 4. Por otro lado, 210 (79%) encuestados expresaron que recibieron preparación para las pruebas Saber 11, 54 estudiantes (21%) opinaron que no fueron preparados. La preparación consiste en responder cuestionarios en época previa a la presentación del examen; 182 (69%) consideran que este entrenamiento les sirvió, mientras que 82 (31%) opinan que no produjo ningún efecto sobre sus resultados. Por otra parte, 142 estudiantes (54%) opinan que los exámenes que hacen los profesores son similares a las pruebas Saber, y 123 (46%) dicen que son diferentes. De igual manera, 177 (67%) creen que las evaluaciones que hacen los profesores sobre lo que les enseñan contribuyen a mejorar el desempeño de la prueba Saber 11, frente a 88 (33%) que opinan que no hay contribución. Finalmente, 176 estudiantes (66%) creen que la enseñanza del colegio aporta al mejoramiento de los resultados del examen, frente a 88 (33%) que consideran que no.

IV. Discusión y conclusiones

Los datos evidencian el deseo de los profesores de Ciencias Sociales de mejorar los resultados de las pruebas Saber de sus estudiantes, pero, además, que la principal acción para hacerlo es contestar cuestionarios que encuentran similares a dichas pruebas en la época previa a su presentación. Tendríamos que preguntarnos qué tipo de aprendizajes promueve el diligenciamiento de cuestionarios si los resultados evidencian que Ciencias Sociales es el área con más bajo rendimiento de las áreas evaluadas; asimismo, qué implicaciones tiene que los estudiantes de último grado creen que están aprendiendo cuando sus resultados evidencian que el área es la de menor rendimiento.

Según la tipología establecida por Wolfe (2006), las pruebas Saber 11 son un examen de evaluación y certificación, producto de la medición de conocimientos y habilidades en las que el reporte de datos se extiende hasta el individuo. Es también una prueba de alto riesgo (Ravela, 2006a), en la medida en que es un examen de acreditación que presentan los estudiantes al finalizar los seis años de educación secundaria, pues sus resultados son utilizados para la selección de los alumnos que desean ingresar a la educación superior, así como para medir la calidad de los colegios, otorgar premios, becas e incentivos económicos. Estas pruebas también tienen una dimensión diagnóstica porque los reportes derivados de ellas brindan información pertinente para el diseño e implementación de planes de mejoramiento. En consecuencia, las pruebas Saber muestran en qué grado el sistema educativo colombiano logra los objetivos de formación planteados para la enseñanza básica, media y superior. En el caso de Ciencias Sociales, es el área de más bajo promedio de todas las evaluadas, según se evidencia en los puntajes obtenidos por los estudiantes en la asignatura (ver figura 3).

Los resultados de las tres instituciones muestran una tendencia en el crecimiento de los puntajes globales de las instituciones educativas públicas de Ibagué; este aumento sería más acelerado en los colegios de la mitad de la lista hacia abajo, y menos acelerado en los colegios de los primeros puestos. En el caso de la Institución I, sube del 5º al 3er puesto con un aumento de 1.3, mientras la Institución II, ubicada, en la mitad de lista, a pesar de aumentar 1.8 baja del puesto 11 al 19, y con un aumento de 2.4, desciende del puesto 25 al 35. Esto indicaría también que, aunque los tres colegios participantes en el estudio subieron sus puntajes, otros tuvieron aumentos más considerables; si esta tendencia se mantuviera en los próximos años, estaríamos ante una mejora significativa de los aprendizajes que son evaluados en las pruebas Saber 11. Lo ideal sería que estas instituciones de la parte final de la lista que muestren mejoras, recibieran acompañamiento y seguimiento para que esta ruta sea constante y permanezca en el tiempo.

La diferencia de los puntajes entre la jornada de la mañana y la de la tarde de dos de las instituciones del estudio nos pone frente a la reflexión sobre por qué en una misma institución los estudiantes matutinos

tienen mejores resultados que los vespertinos. En un trabajo acerca de la relación entre mañana y tarde y la ansiedad en escolares, Díaz-Morales, Collado-Mateo, Escribano y Delgado (2013) observaron una muestra de 638 adolescentes entre los 12 y 14 años de edad, y determinaron que quienes estudiaban en jornada vespertina tenían mayor ansiedad, posiblemente debido a la disparidad entre el desarrollo evolutivo y las dinámicas sociales. Los planteamientos de Ridao y Gil (2002) concluyeron que en el ámbito educativo la determinación de la jornada escolar del alumno está estrechamente relacionada con las funciones pedagógicas de la escuela. Otras variables más relacionadas con los estilos pedagógicos, las interacciones en el aula o la propia distribución del tiempo y de las actividades dentro de la jornada serían responsables más directas de los resultados que se obtengan.

La inclusión de las opiniones de los alumnos sobre el entrenamiento que reciben para las pruebas Saber 11 y la evaluación de la clase de Ciencias Sociales indica que si bien es importante la lectura de los resultados del examen, es necesario también interpretarlos a la luz de otras variables, como las dimensiones socioeconómicas, las características del profesorado y de la organización escolar y el componente pedagógico y didáctico, pues además de comprensiones más complejas, estaríamos ante información útil para el diseño, ejecución y evaluación de planes de mejoramiento escolar. Según Ravela (2006a) el reto es superar la tendencia que han tenido la mayoría de los informes nacionales en América Latina, de ser extremadamente descriptivos y limitarse a relatar los datos emergentes, sin analizar conexiones con otro tipo de información recogida en la misma prueba o proveniente de otras investigaciones.

Las respuestas de los estudiantes a las preguntas sobre la evaluación de la asignatura de Ciencias Sociales y su impacto en los resultados de las pruebas Saber 11 y sobre el uso de cuestionarios como parte del entrenamiento para presentarlas (ver figura 6) ponen sobre la mesa una vez más la reflexión sobre si se están priorizando los aprendizajes de los alumnos o la obtención de mejores resultados en una evaluación. El asunto reviste la mayor importancia por cuanto desde el punto de autores como Ravela (2006b), los críticos del énfasis creciente en los resultados de las pruebas argumentan que otorgar excesiva importancia sobre algunas competencias y algunas habilidades distorsiona la educación, por cuanto otros objetivos primordiales en educación primaria y secundaria son dejados de lado al no evaluarse en los exámenes.

La gran relevancia que se le da a los resultados de los exámenes estandarizados y no tanto a los procesos educativos conducentes a mejores aprendizajes puede llevar a que se enseñen las cosas específicas y necesarias para pasar las pruebas, en lugar de dominios de conocimientos más generales; en el peor de los casos, la enseñanza se puede orientar a contestar mecánicamente un examen y a identificar claves para detectar la respuesta correcta. En palabras de Ravela (2006b), el puntaje de la prueba y el grado de comprensión no son lo mismo, se requiere de una generalización y es la generalización, no la prueba en sí misma, lo que importa; cuando se enseña a los estudiantes a contestar la prueba indicándoles la manera de identificar respuestas correctas e incorrectas se pierde la validez de la inferencia sobre los verdaderos logros alcanzados por los estudiantes.

Con el artículo se espera llamar la atención sobre la estrecha relación entre la enseñanza y los resultados de las pruebas que evalúan competencias, esta discusión también ha sido resaltada por Palacios (2012), Gómez, Miralles y Molina (2015) y Trigueros, Ortuño y Molina (2015), quienes coinciden en la persistencia de metodologías que privilegian los contenidos, la memorización y no propician el desarrollo de un mínimo de destrezas en el área de Ciencias Sociales. Para estos autores, hay todavía una fuerte presencia de actividades en las que se solicita a los estudiantes el aprendizaje de fechas, datos y hechos sin ninguna comprensión de los mismos, mientras la prueba verifica la apropiación de conceptos y de habilidades cognitivas, procedimentales y actitudinales.

Para Wolff (2006), las evaluaciones no son sólo un medio para supervisar el avance en el logro de los objetivos, también es importante evaluar los factores ligados a la construcción de los aprendizajes como los materiales escolares, la capacitación docente, los coeficientes profesor-alumno, las interacciones y la pedagogía en el salón de clases. Por su parte, Ravela (2006a) hace énfasis en que se requiere pasar de la recolección de información cuantitativa sobre los resultados de los exámenes al conocimiento

compreensivo “de lo que realmente se enseña en las escuelas” (p. 103). Además, subraya la necesidad de avanzar en el desarrollo de una metodología para el diseño de los instrumentos, que dé relevancia a los factores escolares como los tipos de escuela, las prácticas de enseñanza eficaces y otros aspectos sustanciales en el marco de un operativo nacional de evaluación, que puedan ser incluidos en la toma de decisiones tanto en la política educativa como dentro de los establecimientos.

Un ejercicio que puede ser valioso para establecer tendencias y explicaciones más complejas de la situación de la educación colombiana es la lectura de los resultados de la prueba Saber 11 en las áreas evaluadas, en diálogo con otros estudios que tratan de explicar factores de diferente índole que influyen en dichos resultados, es el caso del trabajo de Abadía y Bernal (2016), en el que demuestran que hay puntajes globales más sobresalientes en los niños que en las niñas, y que aunque generalmente éstas sobresalen en lectura crítica, los niños les llevan ventajas en matemáticas y ciencias, provocando que éstas tengan menor oportunidades de ingresar a carreras de ciencias, tecnología, ingeniería y matemáticas; o el trabajo de Rodríguez y Tascón (2017), quienes identificaron que la variación en el número de estudiantes de un salón de clases no tiene un peso significativo en los resultados obtenidos en las pruebas Saber 11.

El logro de un mayor impacto en los resultados de las pruebas estandarizadas parece estar relacionado con la generación de espacios de discusión entre los productores de información, los docentes y directivos docentes, los planificadores de las políticas educativas y los capacitadores de docentes (Cueto, 2005). Para este autor, los informes que presentan los resultados de los exámenes son interesantes, pero mucha de la información producida no llega a los destinatarios finales (docentes y padres de familia), y cuando llega, a menudo no es discutida ni interpretada. En esta misma línea, Valverde (2006) y Esquivel (2006) consideran que quienes diseñan las pruebas tienen la responsabilidad de reportar claramente su utilidad y los usuarios de los resultados tienen la responsabilidad de aprovecharlos según los criterios de validez que tienen.

Referencias

- Abadía, L. K. y Bernal, G. (2016). Brechas de género en el rendimiento escolar a lo largo de la distribución de puntajes: evidencia pruebas Saber 11. *Universitas Económica*, 16(9), 1-34. Recuperado de https://cea.javeriana.edu.co/documents/153049/2786252/Vol.16_N9_SEP_2016.pdf/4898df5a-0e0a-47d8-ae10-f98627740d5b
- Barrado, C., Gallego, I. y Valero-García, M. (1999). *Usemos las encuestas a los alumnos para mejorar nuestra docencia*. Barcelona: Universidad Politécnica de Cataluña.
- Díaz-Morales, J. F., Collado-Mateo, M. J., Escribano, C. y Delgado, P. (2013). Matutinidad-vespertinidad y ansiedad rasgo en adolescentes. *Anales De Psicología*, 29(1), 90-93. doi:[10.6018/analesps.29.1.138821](https://doi.org/10.6018/analesps.29.1.138821)
- Cueto, S. (2005). *Información empírica y desarrollo de políticas educativas en América Latina*. Washington: Grupo de Análisis para el Desarrollo [Grade] / Diálogo Regional de Educación del Banco Interamericano de Desarrollo. Recuperado de <http://www.grade.org.pe/publicaciones/758-informacion-empirica-y-desarrollo-de-politicas-educativas-en-america-latina/>
- Departamento Administrativo Nacional de Estadística (2011). *Encuesta de convivencia escolar y circunstancias que la afectan [Ececa], para estudiantes de 5 a 11, año 2011*. Bogotá: Autor.
- Departamento Administrativo Nacional de Estadística (2017). *Estadísticas de Educación*. Bogotá: Autor.
- Esquivel, J. M. (2006). El diseño de las pruebas para medir el logro académico: ¿referencia a normas o a criterios? En P. Arregui (Ed.), *Sobre estándares y evaluaciones en América Latina* (pp. 82-97). Santiago de Chile: PREAL. Recuperado de <http://repositorio.grade.org.pe/handle/GRADE/399>

Ferrer, G. (2005). *Estado de situación de los sistemas nacionales de evaluación de logros de aprendizaje en América Latina*. Santiago de Chile: PREAL.

Ferrer, G. y Arregui, P. (2006). Las pruebas internacionales de aprendizaje en América Latina y su impacto en la calidad de la educación: Criterio para guiar futuras aplicaciones. En P. Arregui (Ed.), *Sobre estándares y evaluaciones en América Latina* (pp. 249-294). Santiago de Chile: Programa de Promoción de la Reforma Educativa de América Latina (PREAL).

Gómez, C., Miralles, P. y Molina S. (2015). Evaluación, competencias históricas y educación ciudadana. *Estudios Sociales*, 52, 52-68. Recuperado de <https://revistas.uniandes.edu.co/doi/pdf/10.7440/res52.2015.01>

Locatelli, I. y Couto, A. (2001). *Diseminación de los resultados de las evaluaciones educacionales a gran escala: la experiencia de Brasil*. Brasil: Ministerio de Educación-Instituto Nacional de Estudio e Investigaciones Educativas.

Ministerio de Educación Nacional de Colombia. (2006). *Estándares Básicos de Competencias en Lenguaje, Matemáticas, Ciencias y Ciudadanas. Guía sobre lo que los estudiantes deben saber y saber hacer con lo que aprenden*. Bogotá: Autor.

Ministerio de Educación Nacional e Instituto Colombiano para la Evaluación de la Educación [MEN-ICFES]. (2013). *Sistema de evaluación y estandarización. Alineación del examen Saber 11*. Bogotá: Autor.

Ministerio de Educación Nacional e Instituto Colombiano para la Evaluación de la Educación [MEN-ICFES]. (2015). *Guía de interpretación y uso de resultados de las pruebas Saber 11. Versión 1*. Bogotá: Autor.

Ministerio de Educación Nacional e Instituto Colombiano para la Evaluación de la Educación [icfes] (2016). Reporte de resultados por aplicación del examen Saber 11 para establecimientos educativos. Bogotá: Autor.

Ministerio de Educación Nacional e Instituto Colombiano para la Evaluación de la Educación [MEN-ICFES]. (2016). *Clasificación de planteles educativos*. Recuperado de <https://www.icfes.gov.co/clasificacion-de-planteles>

Montoya, S., Perusia J. C. y Vera, A. (2001). *Evaluación de la calidad educativa: de los sistemas centrales al aula. Estudio del impacto de las políticas de devolución de información personalizada a las escuelas, en tres provincias de Argentina*. Buenos Aires: Centro de Investigaciones para la Equidad y Calidad Educativa.

Palacios, N. (2012). Los estándares de competencias ciudadanas en Colombia en los libros de texto. *Scripta Nova. Revista Electrónica de Geografía y Ciencias Sociales*, 16(418), 1- 6. Recuperado de <http://www.ub.edu/geocrit/sn/sn-418/sn-418-32.htm>

Perassi, Z. (2008). *La evaluación en educación: un campo de controversias*. Argentina: Facultad de Ciencias Humanas de la Universidad de San Luis y Laboratorio de Alternativas Educativas. Recuperado de http://lae.unsl.edu.ar/Ediciones/Libros_Electronicos/Libro_La_Eval_En_Educacion.pdf

Ravela, P. (2006a). ¿Cómo presentan sus resultados los sistemas nacionales de evaluación educativa en América Latina? En P. Arregui (Ed.), *Sobre estándares y evaluaciones en América Latina* (pp. 181-248). Santiago de Chile: PREAL.

Ravela, P. (2006b). Los resultados de las evaluaciones educativas en la prensa. En P. Arregui (Ed.), *Sobre estándares y evaluaciones en América Latina* (pp. 295-384). Santiago de Chile: PREAL.

Ravela, P., Wolfe, R., Valverde, G. y Esquivel, J. (2006). Los próximos pasos: ¿Cómo avanzar en la evaluación de aprendizajes en América Latina? En P. Arregui (Ed.), *Sobre estándares y evaluaciones en*

América Latina (53- 60). Santiago de Chile: PREAL. Recuperado de <http://repositorio.grade.org.pe/handle/GRADE/399>

Ridao, I. y Gil, J. (2002). La jornada escolar y el rendimiento de los alumnos. *Educación*, 327, 141-156.

Rodríguez, M. y Tascón, D. (2017). Impacto del número de estudiantes de un mismo salón sobre las pruebas de Estado en Colombia. *Vniversitas Económica*, 17(2), 1-33.

Rojas, R. (2014). *Competencias en lenguaje: relaciones entre la escuela y las pruebas de Estado. Estudios sobre la calidad de la educación en Colombia*. Bogotá: MEN-ICFES.

Tiana, A. (2003). ¿Qué pretendemos evaluar, qué evaluamos y qué conclusiones podemos extraer de la evaluación? En: *Evaluar las evaluaciones. Una mirada política acerca de las evaluaciones de la calidad educativa*. Buenos Aires: IPE-UNESCO.

Trigueros, C., Ortuño, J. y Molina, S. (2015). La percepción del alumnado de educación secundaria sobre el desarrollo del pensamiento histórico en el proceso evaluador. *Estudios Sociales*, 52, 69-86. Recuperado de doi:[10.7440/res52.2015.05](https://doi.org/10.7440/res52.2015.05)

Vallejo, G. y Escudero, J. R. (1999). Cuestionarios para evaluar las actitudes de los estudiantes de ESO hacia las matemáticas. *Aula Abierta*, 74, 1-17.

Valverde, G. (2006). La interpretación justificada y el uso apropiado de los resultados de las mediaciones. En P. Arregui (ed.), *Sobre estándares y evaluaciones en América Latina* (70-81). Santiago de Chile: Preal.

Wolfe, R. (2006). El dilema de la granularidad en el diseño del sistema de evaluación: cobertura curricular vs. cobertura poblacional. En P. Arregui (Ed.), *Sobre estándares y evaluaciones en América Latina* (61-69). Santiago de Chile: PREAL.

Wolff, L. (2006). Las evaluaciones educacionales en América Latina: avance actual y futuros desafíos. En P. Arregui (ed.). *Sobre estándares y evaluaciones en América Latina* (13-52). Santiago de Chile: PREAL.