

SABER

AL > DETALLE

EDICIÓN

04

Bogotá D.C.

Abril de 2019

ISSN: 2590-4663

Publicación trimestral



La educación
es de todos

Mineducación

Instituto Colombiano para la Evaluación de la Educación, ICFES

Oficinas: Calle 26 NO. 69-76 Torre 2, piso 15

Edificio Elemento, Bogotá • Colombia

Directora General: María Figueroa Cahnspeyer

Directora de Evaluación: Natalia González Gómez

Subdirectora de Análisis y Divulgación: Ana María Restrepo Sáenz

Subdirección de Estadísticas: Jorge Mario Carrasco Ortiz

Subdirección de Diseño de Instrumentos: Javier Toro Baquero

Coordinación General: Dirección de Evaluación

**¿CÓMO SE CONSTRUYE
EL ÍNDICE DE NIVEL
SOCIOECONÓMICO (INSE)
EN EL CONTEXTO DE LAS
PRUEBAS SABER?**

¿CÓMO SE CONSTRUYE EL ÍNDICE DE NIVEL SOCIOECONÓMICO (INSE) EN EL CONTEXTO DE LAS PRUEBAS SABER?

En cada una de las pruebas Saber se incluyen cuestionarios adicionales que permiten recolectar información acerca del entorno de los estudiantes, tales como antecedentes escolares, competencias socioemocionales y características socioeconómicas y culturales. Estas variables, conocidas como factores asociados al aprendizaje, son relevantes en el estudio de la calidad de la educación, ya que tienen influencia sobre el logro educativo. En particular, se ha documentado ampliamente¹ que el nivel socioeconómico tiene una alta incidencia sobre el desempeño académico. Por lo tanto, para el Icfes es indispensable contar con un indicador que logre capturar las dimensiones que componen el nivel socioeconómico de cada estudiante. Dentro de la

población evaluada en cada prueba, es probable que existan grupos de estudiantes que compartan características comunes que no necesariamente se logren identificar a partir del estrato socioeconómico, ya que dicha variable busca hacer una clasificación según inmueble residencial y no necesariamente apunta al acceso que tiene un estudiante de bienes y servicios. Así, se construye el índice de nivel socioeconómico (INSE) a partir de la información auxiliar recolectada, que permite caracterizar a la población de una manera integral. Como se verá más adelante, los descriptores asociados a cada nivel socioeconómico permiten tener un panorama más profundo sobre la pertenencia de los estudiantes a un nivel particular y permite ver diferencias entre niveles.

¿Qué es el INSE?

El nivel o estatus socioeconómico es un constructo teórico que pretende representar una posición económica o un estatus socioeconómico. Aunque no existe un consenso de lo que precisamente representa, la medida más empleada se centra en una definición basada en capital propuesta por Coleman (1988): capital físico (recursos materiales), capital humano (recursos humanos como educación) o capital social (recursos sociales) que están ligados al bienestar de los individuos (Bradley & Corwyn, 2002). Desde las ciencias sociales, el estatus socioeconómico hace referencia al acceso a recursos colectivamente deseables de bienes, dinero, poder, redes sociales, salud, tiempo libre y oportunidades educativas (Oakes & Rossi, 2003). Desde la psicología, abarca el ingreso, la educación, la seguridad financiera y las percepciones subjetivas sobre el estatus social y clases sociales (APA, 2007). También se define como una medida total que combina la parte económica y sociológica de la preparación laboral de una persona y de la posición

económica y social individual o familiar en relación a otras personas (Vera-Romero & Vera-Romero, 2013). Como se observa, no hay una única forma de medir el nivel socioeconómico.

Si seguimos la definición más empleada que se sustenta en un enfoque material, el nivel socioeconómico es un rasgo latente² que debe resumir información de al menos tres dimensiones (1) nivel educativo de los padres, (2) ocupación de los padres y (3) el ingreso familiar (NCES, 2012). En los cuestionarios sociodemográficos que acompañan las pruebas Saber se incluyen dichas dimensiones. El nivel educativo de los padres se captura con el último nivel educativo alcanzado completo o incompleto de la madre y del padre. En cuanto a la ocupación de los padres, se contempla una serie de oficios que incluye el trabajo en el hogar y la ausencia de trabajo. Por último, se emplea la dotación del hogar como proxy del ingreso familiar, medido a través del consumo de bienes y acceso a servicios.

¹ White (1982) y Sirin (2005) presentan meta-análisis que han profundizado en el estudio de las relaciones entre el estatus socioeconómico y el desempeño académico, y encuentran que existe una relación positiva entre ellos.

² Es un constructo que no puede ser observado directamente.

¿Cuál es la metodología empleada para el cálculo del INSE?

Desde 2009 hasta 2012 el INSE era calculado con base en la metodología de análisis multivariante que emplea el Departamento Nacional de Planeación (DNP), la cual consiste analizar las relaciones existentes entre múltiples variables que apuntan a medir el nivel socioeconómico. A partir de 2012 y hasta la fecha, en las pruebas Saber 359, Saber 11, Saber Pro y Saber TyT se emplea la metodología de teoría respuesta al ítem (TRI) para dicho cálculo. Una gran ventaja de emplear TRI respecto al análisis multivariante es que es posible construir una escala histórica que mantiene la escala de medición entre periodos, mientras que con el análisis multivariante la escala puede variar ligeramente. En ese sentido, se garantiza comparabilidad entre periodos.

La TRI parte del supuesto que un rasgo latente puede medirse mediante las respuestas a los ítems o preguntas (DeMars, 2010). De esta manera, una variable no observada, como el índice socioeconómico, puede ser medida a través de la respuesta a ítems propuestos. Como se aborda en la primera edición de Saber al Detalle, la TRI modela en términos probabilísticos la relación que existe entre las respuestas que un individuo le da a un conjunto de ítems y una variable que no se puede observar directamente, como el INSE. Para ello, es necesario validar que las variables empleadas para la medición del nivel socioeconómico sean las adecuadas, lo cual se analiza a través de comportamiento psicométrico y estadístico de cada ítem de medición del INSE. Posteriormente, se estima el modelo TRI.

¿Cómo se realiza el análisis de ítems?

Este análisis permite determinar si las variables apuntan a explicar un único atributo, que en este caso es el nivel socioeconómico de los evaluados. Lo que se espera lograr con este proceso es generar evidencias estadísticas que permitan mejorar el proceso de selección de ítems y emplear aquellos ítems que presenten la mayor relación posible con el nivel socioeconómico de los estudiantes. Por tanto, se obtiene una mejor medición del INSE para el periodo de observación y para futuras mediciones.

Se emplea un conjunto de variables extenso para la medición del nivel socioeconómico, que al ser consideradas en su totalidad pueden llegar a aportar marginalmente poco a la medición. En ese sentido, es posible sintetizar la información en un conjunto reducido de componentes, mientras se preserva la mayor cantidad de información a través del análisis de componentes principales (ACP). Este análisis permite (1) estudiar las relaciones que hay entre variables, (2) verificar que las variables seleccionadas para la medición apunten a medir una misma dimensión, (3) observar el signo de la relación y (4) la comunalidad que tienen³.

A manera de ejemplo, si se tienen 10 variables para la medición del INSE, después de un ACP la información puede ser resumida en 2 dimensiones que explican el 80% de variación del nivel socioeconómico, y tan solo sea necesario

descartar 3 variables que apunten hacia otra dimensión. Tales dimensiones pueden estar asociadas positivamente con el nivel socioeconómico, y por tanto, aumentos en cada dimensión apuntan a niveles socioeconómicos más altos. Cuando se indaga por el estrato socioeconómico del estudiante, este se emplea como variable de control del índice construido, ya que es de esperar que entre más alto sea el estrato al que pertenece el estudiante lo sea el puntaje del INSE. De esta manera, el estrato socioeconómico permite validar el índice; sin embargo, no se tiene en cuenta en el ACP. Adicionalmente, se realiza una validez del constructo, lo cual busca confirmar que las inferencias que se contemplan sobre el constructo en teoría se vean reflejadas en los datos.

Como se menciona anteriormente, contar con la misma métrica permite caracterizar adecuadamente el nivel socioeconómico al que pertenecen los estudiantes, por lo cual es necesario comparar entre periodos el comportamiento de los ítems que se emplean para medir el INSE. Recordemos que la TRI permite estimar el nivel socioeconómico, que es nuestro constructo de interés, con base en las respuestas dadas a los ítems y las características o parámetros de los ítems. Así, la función que representa la probabilidad de elegir una respuesta en términos del nivel socioeconómico y los parámetros del ítem resulta ser la curva característica del ítem⁴, con la cual podemos analizar las calibraciones de los parámetros de los

³ La comunalidad es uno de los términos más clásicos del análisis factorial, que se refiere a qué tan dispersos están los factores comunes. A mayor porcentaje de variabilidad entre los factores, menos comunalidad hay; y a menor porcentaje de variabilidad entre factores, mayor comunalidad.

⁴ Para profundizar sobre TRI y las curvas características del ítem refiérase al documento 1 “¿Cómo se generan los puntajes en las pruebas Saber del ICES?”.

ítems entre periodos, y por tanto, analizar si existe un comportamiento diferencial del ítem (DIF, por sus siglas en inglés). En este proceso se determina si los ítems o preguntas han presentado cambios significativos en su comportamiento, ya sea por modificaciones en la estructura del ítem o en las características de la población observada. Así, se busca comparar la probabilidad que tiene un estudiante de cierto nivel socioeconómico de contestar alguno de los ítems en el periodo actual (grupo referencia), respecto a un estudiante de un periodo previo (grupo focal) que pertenece al mismo nivel socioeconómico frente al mismo ítem.

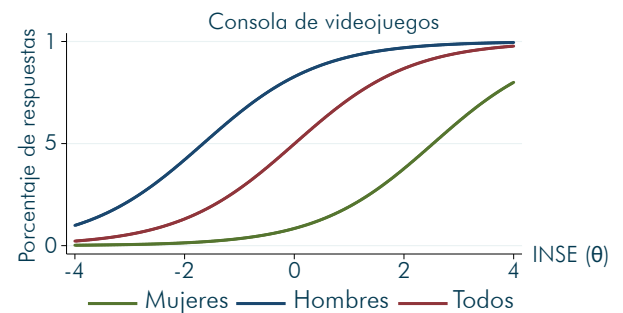
En el caso de una variable con dos valores de respuesta posible (variable dicótoma), como por ejemplo la posesión de moto, se analiza la probabilidad que tienen estudiantes de cierto nivel socioeconómico de poseer una moto entre un periodo y otro. Por otro lado, frente a una variable con más de dos valores de respuesta posible (variable politómica), como por ejemplo nivel de educación del padre, se observa la probabilidad de que los padres de estudiantes pertenecientes a ciertos nivel socioeconómico se encuentren dentro de alguna de las categorías de la variable entre un periodo y otro, como por ejemplo “Nivel de educación profesional completa”.

Si se determina en el análisis DIF que el ítem no presenta cambios estadísticamente significativos entre periodos, se asignan los valores de la estimación anterior, lo cual se conoce como anclaje⁵. Por el contrario, si se determina en el análisis de DIF que sí hay cambios significativos de los ítems entre periodos se procede a liberarlo, lo que conlleva a estimar de nuevo los valores de los parámetros de todos los ítems de manera individual. Si existe un nuevo ítem que se quiera incluir en la medición del INSE se deben estimar los valores de los parámetros del ítem, ya que no se cuentan con estimaciones previas que nos den información sobre su comportamiento.

Luego, con el fin de garantizar comparabilidad entre aplicaciones, las estimaciones de los parámetros de los ítems deben ser fijados a los ítems de anclaje.

Veámonos como se ve a través de curvas características de los ítems un análisis del comportamiento de las variables por características de interés del tipo descriptivo. Para ilustrar lo anterior, se muestra el comportamiento del ítem “consola de videojuegos”.

Figura 1. Comportamiento del ítem



Fuente: ICFES, 2019

Como se observa en figura 1, la variable presenta buena discriminación para definir el nivel socioeconómico, ya que la curva característica del ítem “Todos” muestra que para niveles bajos de INSE la probabilidad de tener consola de videojuegos es cercana a cero, y a medida que aumenta el INSE también lo hace la probabilidad de poseer este artículo en su hogar. Sin embargo, se detecta que al tener en cuenta el sexo, la probabilidad de tener consola de videojuegos para un INSE alto es mayor para los hombres que para las mujeres, ya que al comparar las tres curvas características del ítem (Todos, Hombres y Mujeres) la curva de las mujeres se encuentra desplazada a la derecha.

¿Cómo se realiza la estimación del modelo TRI?

Una vez validado que los ítems apunten a la medición del nivel socioeconómico, y por tanto que haya unidimensionalidad, se procede a la estimación del modelo con la metodología de TRI para calcular el puntaje del INSE. Se estiman los parámetros de los ítems de cada uno de ellos para determinar el nivel socioeconómico de los estudiantes. Hay un símil entre el cálculo del INSE y la clasificación de las pruebas Saber: dados unos valores de los parámetros de los

ítems, para cada nivel del INSE, que en este caso sería nuestro rasgo latente (θ), hay una probabilidad de elección (o acierto) de una categoría de respuesta de cada ítem.

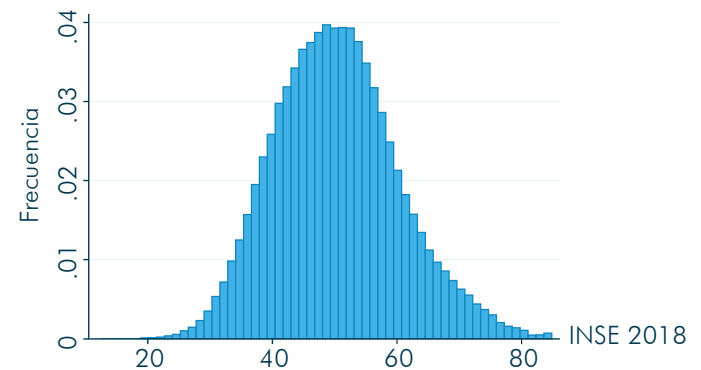
Si las variables son dicótomas, se ajusta un modelo logístico de 2 parámetros (2PL). En el caso de variables politómicas, es decir que tienen varias opciones de respuesta, es posible realizar dos procedimientos. Si las opciones de respuesta se

⁵ Para profundizar el proceso de comparabilidad de los resultados refiérase a la Edición 3 del Boletín Saber al Detalle. ¿Qué garantiza la comparabilidad de los resultados en las pruebas saber realizadas por el Icfes?

pueden ordenar, se realizan asociaciones para cada categoría de la variable con el nivel socioeconómico a través de un modelo de respuesta graduada. Por ejemplo, qué tanto se asocia el nivel educativo alcanzado por la madre “Posgrado” a un puntaje de INSE alto. Pero, si las opciones de respuesta simplemente definen un atributo y no importa el orden, se ajusta un modelo de respuesta nominal.

Teniendo en cuenta que bajo TRI la escala del trazo latente está en el rango de $-\infty$ a ∞ , y que con una probabilidad del 99% se encuentra entre -4 y 4, se reescala el puntaje. De forma tal que el INSE tiene un rango de [0 a 100], está centrado en 50 y se dispersa 10 unidades alrededor de la media. Para ver lo anterior, se presenta en la figura 2 un histograma del puntaje del INSE para el periodo 2018.

Figura 2. Distribución del INSE



¿Cuál es la diferencia entre INSE y el NSE?

El puntaje del INSE se emplea como insumo para el cálculo del nivel socioeconómico categórico (NSE). El NSE se usa para caracterizar la población que presenta las pruebas Saber y se genera como reporte tanto para los establecimientos educativos como para los evaluados. Para su cálculo se tiene en cuenta los dos últimos periodos registrados del INSE y se agrega el periodo actual de estudio, donde se calcula el promedio por año del INSE con su respectiva desviación para realizar comparaciones.

El NSE se enmarca en una escala de NSE 1 hasta NSE 4, donde el primer nivel hace referencia a estudiantes pertenecientes a niveles socioeconómicos bajos y se incrementa hasta el cuarto nivel, que corresponde a estudiantes pertenecientes a niveles socioeconómicos altos. Esta categorización permite tener en cuenta no solo el nivel de ingresos, sino también posesión de bienes, acceso a servicios y educación del núcleo familiar, lo cual brinda una visión completa del hogar del estudiante. En la tabla 1 presentan los puntos de corte del puntaje del INSE para definir qué nivel del NSE caracteriza a cada estudiante.

Tabla 2. Caracterización por NSE

Tabla de frecuencias - Descriptores Socioeconómicos

Variable de caracterización		Niveles NSE categórico			
		1	2	3	4
Internet	(No)	X	X		
Computador	(No)	X	X		
Horno microondas o a Gas	(No)	X			
Lavadora	(No)	X			
Educación de la madre:	Primaria incompleta	X			
Lavadora	(Sí)		X		
Computador	(Sí)		X		X
Servicio de televisión	(Sí)		X	X	
Educación de madre:	Secundaria completa			X	
Automóvil particular	(No)			X	
Horno microondas o a Gas	(Sí)			X	
Automóvil particular	(Sí)			X	X
Educación de la madre:	Profesional completa				X
Internet	(Sí)				X
Consola de videojuegos	(Sí)				X

Tabla 1. Puntos de corte del INSE por NSE

Nivel socioeconómico	Puntos corte INSE
NSE 1	0 - 41,109
NSE 2	Hasta 51,176
NSE 3	Hasta 64,080
NSE 4	Hasta 100

Fuente: Icfes, 2019

Fuente: Icfes, 2019

De los descriptores anteriores, se presenta la siguiente caracterización:

- Los estudiantes que presentan Saber 11 pertenecientes al NSE1 suelen no acceder servicio de internet ni tener computador, horno microondas ni lavadora. Típicamente la educación de la madre es primaria incompleta.
- Los estudiantes de NSE2 mantienen la ausencia de servicio de internet observada para los estudiantes de NSE1, pero la tenencia de computador es menos marcada. Este nivel socioeconómico se caracteriza principalmente por la tenencia de lavadora y el acceso a servicio de televisión.
- Para los niveles NSE3 y NSE4 se observa una diferencia notoria respecto a NSE1 y NS2. Los estudiantes pertenecientes a NSE3 se caracterizan por tener acceso a

servicios de televisión, como lo presentan los estudiantes en NSE2, pero es característico tener horno microondas y no poseer automóvil. Adicionalmente las madres de los estudiantes presentan típicamente un nivel educativo de secundaria completa.

- Para los estudiantes pertenecientes al NSE4, es característico tener computador, automóvil, servicio de internet y consola de videojuegos. A diferencia del NSE3, las madres de los estudiantes pertenecientes a NSE4 tienen educación profesional completa. Finalmente, los estudiantes enmarcados dentro del NSE4, se caracterizan por la posesión de computador y de automóvil particular en sus hogares, y suelen contar con servicio de internet y poseer una de consola de videojuegos. En lo referente a la educación de los padres, la característica predominante es que la madre suele tener un nivel educativo de carrera profesional terminada como máximo nivel educativo.

Bibliografía

- APA. (2007).** *Report of the APA Task Force on Socioeconomic Status*. Washington DC: American Psychological Association.
- Bradley, R. H., & Corwyn, R. F. (2002).** Socioeconomic Status and Child Development. *Annual Review of Psychology* Vol. 53, 371-399.
- DeMars, C. (2010).** *Item response theory*. New York, New York: Oxford University Press, Inc.
- NCES. (2012).** *Improving the Measurement of Socioeconomic Status for the National Assessment of Education Progress: A Theoretical Foundation*. 2012: National Center for Education Statistics (ED).
- Oakes, J. M., & Rossi, P. H. (2003).** The measurement of SES in health research: current practice and steps toward a new approach. *Social Science & Medicine* 56, 769-784.
- Sirin, S. R. (2005).** Socioeconomic status and academic achievement: a meta-analytic. *Review of Educational Research*, 75(3), 417-453.
- Vera-Romero, O., & Vera-Romero, F. M. (2013).** Evaluación del nivel socioeconómico: presentación de una escala adaptada en una población de de una escala adaptada en una población de Lambayeque. *Revista cuerpo médico HNAAA* 6(1) , 41-45.
- White, K. R. (1982).** The relation between socioeconomic status and academic achievement. *Psychological Bulletin*, 91(3), 461-481.

SABER

AL > DETALLE

**¿CÓMO SE CONSTRUYE
EL ÍNDICE DE NIVEL
SOCIOECONÓMICO (INSE)
EN EL CONTEXTO DE LAS
PRUEBAS SABER?**



La educación
es de todos

Mineducación



@icfescol



ICFES



icfescol



YouTube: ICFES

Instituto Colombiano para la Evaluación de la Educación, ICFES

Oficinas: Calle 26 NO. 69-76 Torre 2, piso 15

Edificio Elemento, Bogotá • Colombia