

# SABER

AL > DETALLE



Educación

EDICIÓN

13

Bogotá D.C.

Septiembre de 2023

ISSN: 2590 - 4663

Publicación Trimestral

**¿CÓMO SE REALIZAN  
LOS MUESTREOS  
PROBABILÍSTICOS  
DE LAS PRUEBAS A  
GRAN ESCALA QUE  
DESARROLLA Y EN  
LAS QUE PARTICIPA EL  
ICFES?**

icfes 

**Presidente de la República**  
Gustavo Francisco Petro Urrego

**Ministra de Educación Nacional**  
Aurora Vergara Figueroa

**Viceministro de Educación**  
**Preescolar, Básica y Media**  
Oscar Sánchez Jaramillo



**Elaboración del documento**  
Ana María Mondragón Moreno  
Astrid Julieth Betancourt Pineda  
Nila Fernanda Amaya Melo  
Jesús Fernando Sánchez Velázquez  
Magda Lorena Duarte Ostos  
Jenny Paola Martínez Fonseca  
Patricia Escudero Montañez  
Mishell Marcela Ramos de la Hoz  
Carlos Arturo Parra Villamil

**Diseño y diagramación**  
Aura Cristina Castro Moya

Bogotá D.C., Septiembre 2023

Todos los derechos de autor reservados ©.

**Directora General**  
Elizabeth Blandón Bermúdez

**Secretaria General**  
Lorena Catalina Ramírez Duque

**Directora de Evaluación**  
Julie Paola Caro Osorio (e)

**Subdirector de Diseño de Instrumentos**  
Rafael Eduardo Benjumea Hoyos

**Subdirectora de Análisis y Divulgación**  
Julie Paola Caro Osorio

**Subdirector de Estadísticas**  
Cristian Fabian Montaña Rincón

**Director de Producción y Operaciones**  
Oscar Orlando Ortega Mantilla

**Director de Tecnología e información**  
Carlos Alberto Durán (e)

**Subdirectora de Producción de Instrumentos**  
Daniela Pérez Otavo

**Subdirector de Aplicación de Instrumentos**  
Oscar Orlando Ortega Mantilla (e)

**Subdirector de Desarrollo de Aplicaciones**  
Armando Alfonso Leyton González

**Jefa Oficina Asesora de**  
**Comunicaciones y Mercadeo**  
Johanna Andrea Contreras Valderrama

**Jefa Oficina Asesora de Gestión de**  
**Proyectos de Investigación**  
Jeantit Camilo Cortés Mora

## TÉRMINOS Y CONDICIONES DE USO PARA LAS PUBLICACIONES Y OBRAS QUE SON PROPIEDAD DEL ICfes

El Instituto Colombiano para la Evaluación de la Educación (Icfes) pone a disposición de la comunidad educativa, y del público en general, de forma gratuita y libre de cualquier cargo, un conjunto de publicaciones disponibles en su portal web [www.icfes.gov.co](http://www.icfes.gov.co). Estos materiales y documentos están normados por la presente política y se encuentran protegidos por derechos de propiedad intelectual y derechos de autor a favor del Icfes. Si tiene conocimiento de alguna utilización contraria a lo establecido en estas condiciones de uso, por favor infórmenos al correo [prensaicfes@icfes.gov.co](mailto:prensaicfes@icfes.gov.co).

Queda prohibido el uso o publicación total o parcial de este material con fines de lucro. Únicamente está autorizado su uso para fines académicos e investigativos. Ninguna persona natural o jurídica, nacional o internacional, podrá vender, distribuir, alquilar reproducir, transformar\*, promocionar o realizar acción alguna con la cual se lucre directa o indirectamente con este material. Esta publicación cuenta con el registro ISBN (International Standard Book Number, o Número Normalizado Internacional para Libros), que facilita la identificación, no sólo de cada título, sino, también de la autoría, la edición, el editor y el país dónde se edita.

---

\*La transformación es la modificación de la obra a través de la creación de adaptaciones, traducciones, compilaciones, actualizaciones, revisiones y, en general, cualquier modificación que se pueda realizar, haciendo que la nueva obra resultante se constituya en una obra derivada protegida por el derecho de autor, con la única diferencia, respecto de las obras originales, que aquellas requieren, para su realización, de la autorización expresa del autor o propietario, para adaptar, traducir, compilar, etc. En este caso, el Icfes prohíbe la transformación de esta publicación. Términos y condiciones de uso para las publicaciones y obras que son propiedad del Icfes.

En todo caso cuando se haga uso parcial o total de los contenidos de esta publicación, el usuario deberá consignar o hacer referencia a los créditos institucionales del Icfes, respetando los derechos de cita. En otras palabras, se podrá hacer uso de esta publicación si dicho uso se contempla en los fines aquí previstos. Es posible, entonces, transcribir pasajes del texto si se cita siempre la fuente de autor. Por supuesto, estas citas no deberían ser excesivas ni frecuentes para que, así, no se considere una reproducción simulada y sustancial que redunde en perjuicio del Icfes.

Asimismo, los logotipos institucionales son marcas registradas y de propiedad exclusiva del Icfes. Por tanto, cuando su uso pueda causar confusión, los terceros no podrán usar las marcas de propiedad del Icfes con signos idénticos o similares respecto a cualquier producto o servicio prestado por esta entidad. En todo caso queda prohibido su uso sin previa autorización expresa por parte del Icfes. La infracción de estos derechos se perseguirá civil y penalmente (en caso de que sea necesario) de acuerdo con las leyes nacionales y tratados internacionales aplicables.

*El Icfes realizará cambios o revisiones periódicas a los presentes términos de uso y los actualizará en esta publicación.*

## ¿CÓMO SE REALIZAN LOS MUESTREOS PROBABILÍSTICOS DE LAS PRUEBAS A GRAN ESCALA QUE DESARROLLA Y EN LAS QUE PARTICIPA EL ICFES?

Con el fin de evaluar la calidad de la educación en Colombia, el Icfes ha desarrollado y participado en diferentes pruebas a gran escala. Para algunas de ellas, se han implementado diseños muestrales que contribuyen a garantizar la generación de información de interés para diferentes niveles de agregación, y promueven la toma de decisiones informada y el desarrollo de diferentes investigaciones acerca de la evaluación de la educación.

Dado lo anterior, este boletín tiene como objetivo abordar algunas generalidades de los diseños de muestreo probabilístico utilizados por el Instituto. Para ello, se iniciará con la definición de las pruebas a gran escala, seguida de una introducción sobre los diseños muestrales aplicados en el ámbito educativo y, por último, se presentarán las generalidades de los diseños de las pruebas en las que ha colaborado el Icfes.

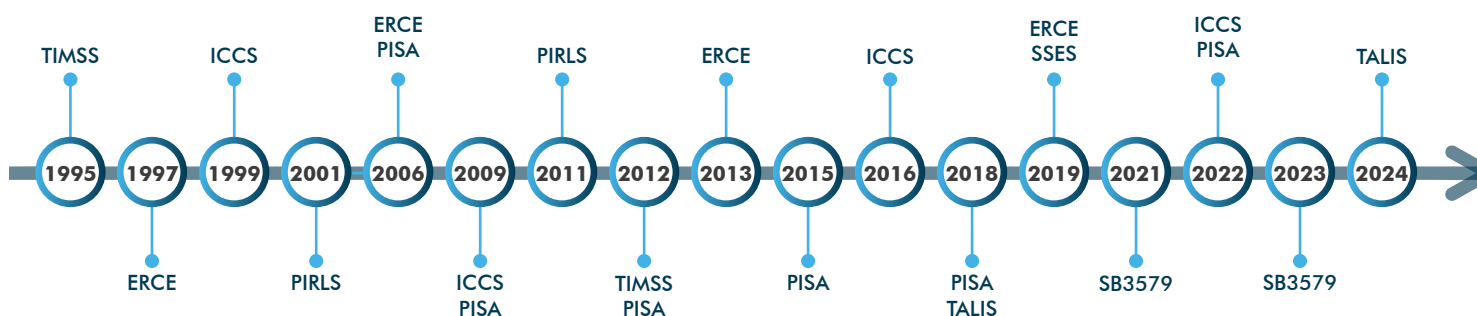
### ¿Qué son pruebas a gran escala?



Las pruebas a gran escala son estudios empíricos basados en instrumentos estandarizados que permiten monitorear los sistemas educativos y, además, recoger información frente a su calidad. Si bien pueden existir otras herramientas para tales fines, estas pruebas son ampliamente utilizadas en el sector educativo. Los resultados obtenidos a partir de ellas son importantes, ya que sirven como soporte o evidencia para la toma de decisiones en materia de políticas educativas a nivel gubernamental o institucional, así como para desarrollar diferentes tipos de investigaciones o simplemente como

medio de información general. El Icfes ha desarrollado pruebas a gran escala a nivel nacional y, además, ha implementado diversas pruebas internacionales en representación de Colombia, lo cual ha permitido analizar el desempeño de los estudiantes y otros actores educativos respecto al de otros países. En la Figura 1 se presenta una línea de tiempo que muestra la trayectoria del Instituto con este tipo de pruebas. Sin embargo, más adelante se proporcionará información respecto al objetivo de cada prueba y el diseño muestral implementado.

FIGURA 1. Línea de tiempo de la aplicación de pruebas a gran escala en Colombia



## ¿Por qué no es necesario implementar siempre un censo?



Las pruebas a gran escala mencionadas en la sección anterior presentan una distinción fundamental en comparación con exámenes como Saber 11°, Saber Pro y Saber TyT. Estas últimas evalúan a toda la población objetivo, mientras que las pruebas a gran escala se llevan a cabo mediante estudios por muestreo. Esta diferencia conlleva una serie de implicaciones significativas. Al realizar pruebas a gran escala a través de un muestreo probabilístico cuidadosamente seleccionado, se busca obtener una representación precisa y confiable de la población objetivo en su conjunto. Esto permite realizar inferencias estadísticas sólidas sobre el desempeño educativo y obtener una imagen más completa y precisa del panorama educacional del país sin necesidad de evaluar a todos los individuos.

Para hablar de las principales diferencias entre un estudio poblacional (censo) y un estudio muestral, es necesario definir cada uno de estos procesos. Un *censo* es un estudio de cada unidad en una población, es decir, una enumeración de cada individuo, mientras que, un *estudio muestral* es un análisis realizado a un subconjunto de la población seleccionado para representar todas las unidades de una población de interés (Australian Bureau of Statistics, 2023a). Cuando hablamos de los valores que describen la totalidad de la población (ej. media poblacional) nos referimos a un *parámetro*, mientras que una *estimación* o *estadística* corresponde a un valor que describe una muestra (ej. media muestral), (Bhandari, 2020). De manera general, las estimaciones se calculan en una muestra para conocer de manera aproximada los parámetros de la población.

**TABLA 1. Comparación entre los conceptos de censo y estudio muestral**

<b>Censo</b>	<b>Estudio por Muestreo</b>
No requiere inferencia, es decir de este se pueden derivar medidas “verdaderas” de la población (no hay error de muestreo).	Con la información de una parte de la población se pueden hacer inferencias de características para toda la población.
Funciona como marco de muestreo <sup>1</sup> para otros estudios.	Las estimaciones e inferencias están sujetas a un error de muestreo.
Permite obtener resultados a un amplio nivel de desagregación.	Dificulta hacer otro tipo de estudios y obtener resultados de otras desagregaciones.
Al tener que observar todos y cada uno de los elementos de la población, resulta ser costoso.	Resulta más económico que llevar a cabo un censo
Logística compleja.	Logística más sencilla.
Los resultados suelen obtenerse en un largo plazo.	Implica más control.
Permiten obtener resultados a nivel de individuo.	Resultados en menor tiempo.
	No están diseñados para generar resultados a nivel de individuo.

**1.** El marco de muestreo se define como el dispositivo que permite identificar, seleccionar y ubicar a todos y cada uno de los objetos pertenecientes a la población objetivo y que participarán en la selección aleatoria. (Gutiérrez, 2009).

El censo es un ejercicio complejo en términos logísticos y la mayoría de las veces costoso, que implica recopilar información detallada sobre una población objetivo<sup>2</sup>. Aunque los censos son herramientas importantes para recopilar datos poblacionales, no se llevan a cabo de forma constante ya que implican un alto gasto en recursos humanos y monetarios, que en consecuencia, dificultan su frecuente actualización. Además, son propensos a tener un error de medición más alto.

De allí, surge la alternativa de realizar mediciones sobre un subconjunto de la población en vez del total, que es lo

que denominamos *estudio por muestreo*. En la Tabla 1 se comparan las ventajas y desventajas de un censo versus un estudio por muestreo.

En el contexto de la investigación educativa, la selección de una muestra brinda varias ventajas respecto al censo. Un proceso de muestreo riguroso permite realizar inferencias precisas e insesgadas frente a la población objetivo que se espera evaluar, por lo que su óptimo desarrollo depende de una definición clara de los objetivos, de una adecuada aplicación y demás pasos que se presentan más adelante.

## ¿Qué es un muestreo probabilístico?



Para garantizar que los resultados de estimación reflejen fielmente el comportamiento de la población, se debe garantizar que el diseño de muestreo sea probabilístico. Existen diversos diseños de muestreo que se aplican a una población dependiendo de los objetivos del estudio, las características de la población objetivo y las estimaciones que se esperan obtener de ella.

En este sentido, un diseño muestral es probabilístico si cumple la condición de que todos los elementos de la población tengan una probabilidad mayor a cero de ser incluidos en la muestra. Por lo tanto, este tipo de diseños se realizan mediante metodologías estadísticas que garantizan lo anterior, y que evitan la selección de una muestra sesgada, al menos desde un punto de vista teórico.

Las condiciones que debe cumplir una muestra para ser probabilística son las siguientes:

1. Se puede definir el conjunto de todas las muestras posibles.
2. A cada muestra posible le corresponde una probabilidad de selección conocida.
3. El proceso de selección garantiza que todos los elementos de la población tengan una posibilidad de ser seleccionados.
4. El mecanismo o algoritmo de selección garantiza que cada muestra de un tamaño dado tenga una posibilidad de ser seleccionada.

Al realizar un diseño probabilístico es posible construir factores de expansión<sup>3</sup>, los cuales permiten que los resultados de un estudio se puedan generalizar a toda la población. De esta manera, cada individuo de la muestra representa a otras unidades no seleccionadas con sus mismas características, y así, la muestra representa la población entera.

**2.** En estadística, se puede definir una población objetivo como la colección completa de todas las unidades que se quieren estudiar. (Gutiérrez, 2009).

**3.** Los factores de expansión se definen como el inverso de la probabilidad de inclusión (Gutiérrez, 2009). Otra definición que se puede encontrar de los factores de expansión es la siguiente: Es un valor que multiplica a cada uno de los elementos seleccionados en la muestra con el fin de obtener la estimación del parámetro en el universo, este factor depende y se construye a partir del método de selección utilizado. (DANE, 2005, p. 11).

Cabe señalar que las pruebas a gran escala muestrales que aplica y en las que participa el Icfes son diseñadas bajo muestreos probabilísticos para poblaciones finitas<sup>4</sup> y más adelante se describirán con mayor detalle.

## ¿Qué diseños de muestreo se aplican en las pruebas a gran escala?



Como se mencionó anteriormente, el tipo de muestreo a aplicar dependerá de la población objetivo y, además, del objetivo del estudio. Dentro de los muestreos más usados en las pruebas a gran escala encontramos el muestreo aleatorio simple, el diseño proporcional al tamaño (PPT) y el muestreo estratificado.

**Muestreo aleatorio simple:** Como bien lo señala su nombre, este es el muestreo más sencillo que podemos aplicar en un estudio. Consiste en seleccionar una muestra de tal manera que se garantice que cada elemento de la población tenga la misma probabilidad de ser seleccionado, lo que implica, que todas las posibles muestras tienen también la misma probabilidad de selección. Algunas de las características más importantes que debe tener la población para aplicar este tipo de muestreo es que todos los elementos que pertenecen a ella sean lo más parecidos posible respecto a la variable de interés, y que, además, puedan ubicarse fácilmente (Australian Bureau of Statistics, 2023b).

**Muestreo proporcional al tamaño:** es uno de los muestreos más utilizados y uno de los más eficientes, ya que posee una menor varianza muestral respecto a otros métodos de selección. En este diseño, se asignan las probabilidades a los elementos de manera que su probabilidad será proporcional a una medida de tamaño (Särndal, 2003); es decir, entre más grande sea esta medida de tamaño en la unidad, mayor será su probabilidad de selección. Por ejemplo, para el diseño

que se aplicó en la prueba Saber 3°, 5°, 7° y 9°, los establecimientos tienen una probabilidad de selección respecto al número de estudiantes matriculados en los grados 3°, 5°, 7° y 9.°, por lo que entre más estudiantes tenga matriculados en los grados antes mencionados, mayor será su probabilidad de selección.

Después, para obtener la muestra, se aplica un método de selección probabilístico dentro de cada estrato. Generalmente en las pruebas diagnósticas de la educación se suele aplicar un muestreo proporcional al tamaño sin reemplazo dentro de cada estrato. Uno de los objetivos de esta estratificación es asegurar la representatividad a través de la inclusión de unidades muestrales de ciertas características en la muestra y, además, lograr la disminución de los errores de muestreo de las estimaciones generadas para estos grupos.

**Muestreo en varias etapas:** este diseño es útil cuando las poblaciones presentan una estructura jerárquica de grupos o conglomerados<sup>5</sup> fácilmente identificable. También, este diseño puede ser útil si estos conglomerados presentan alta variabilidad, gran tamaño, dispersión de tipo geográfico considerable o su medición representa costos logísticos o financieros elevados. Esta técnica consiste en dividir la población o tomar conglomerados previamente definidos, y seleccionar una muestra a partir de estos grupos, lo cual se denominaría primera etapa de muestreo. Luego, cada grupo seleccionado se divide nuevamente en subgrupos y

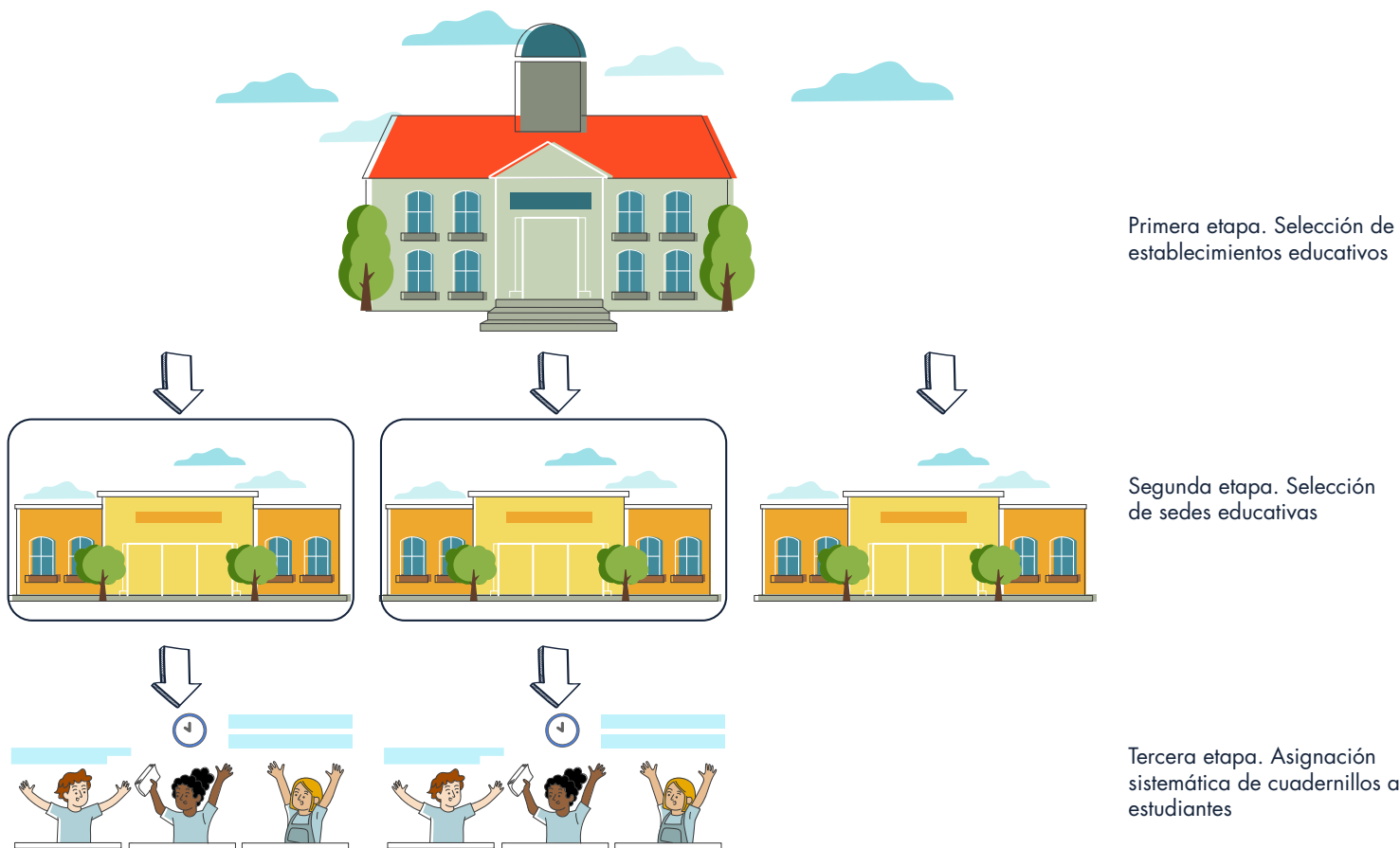
4. Población finita: Es aquella para la cual se puede determinar el número total de elementos que la componen.

5. Conglomerado: agrupación natural de los elementos de una población.

se selecciona otra muestra que se denominaría segunda etapa. Secuencialmente, se sigue este proceso hasta llegar a la unidad de observación final. En cada etapa se puede utilizar cualquiera de los diseños descritos anteriormente.

Por ejemplo, en las pruebas Saber 3°, 5°, 7° y 9° se utiliza este diseño, teniendo en cuenta tres etapas de muestreo (Figura 2).

FIGURA 2. Diseño muestreo aplicado a las pruebas Saber 3°, 5°, 7° y 9°



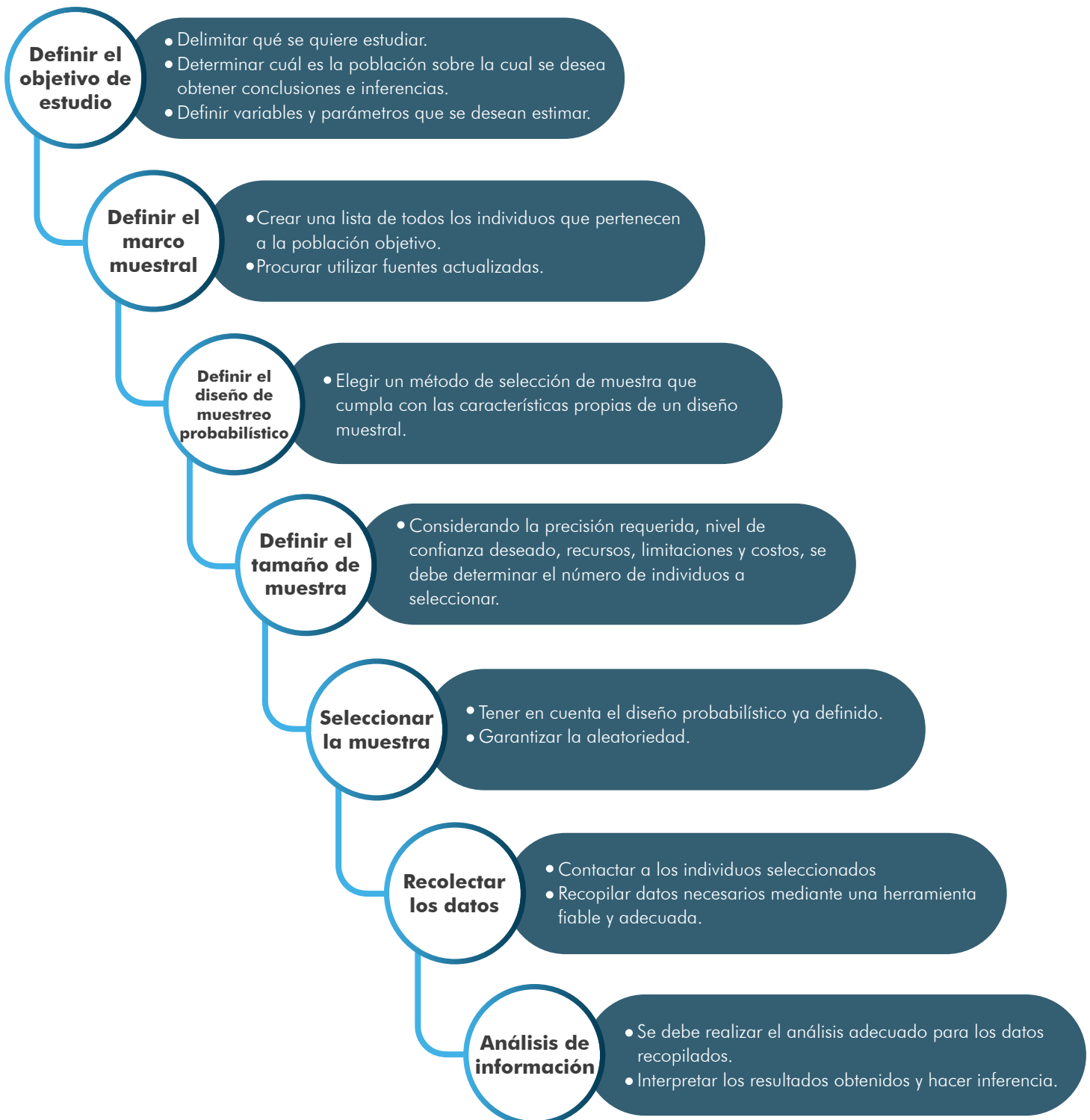
## ¿Cuáles son los pasos para seleccionar una muestra?



Como se mencionó en secciones anteriores, la selección de una muestra depende de diferentes aspectos asociados con la finalidad de la investigación. Sin embargo, en

la Figura 3 se desglosan algunos pasos generales que pueden ser útiles para este proceso.



**FIGURA 3. Pasos para seleccionar una muestra**


Los pasos presentados en la Figura 3 pueden variar dependiendo de las necesidades del estudio. En las pruebas internacionales, por ejemplo, los consorcios pueden establecer previamente algunos elementos

específicos del diseño muestral. Sin embargo, esta estructura puede orientar el proceso para la selección de la muestra.

## ¿Qué resultados obtenemos a partir de la aplicación de estas metodologías?



Es importante resaltar que las pruebas a gran escala corresponden a pruebas diagnósticas que, debido a su metodología, principalmente brindan un diagnóstico a nivel de competencias y cuyo objetivo no es un resultado individual. La aplicación de un diseño muestral dentro del marco de la investigación educativa se realiza, generalmente, con el objetivo de obtener estimativos del comportamiento y frente al desarrollo de competencias en el contexto educativo, así como los cambios que se generan durante el proceso educativo a nivel nacional o regional. Esto brinda información relevante para la implementación de políticas de mejoramiento de la educación. La información generada de la muestra resultante normalmente se utiliza para realizar inferencias sobre la población total objetivo de estudio. Estas generalizaciones pueden ser estimaciones de una o más características asociadas con la población, o pueden estar relacionadas con las estimaciones de la magnitud de la asociación entre las características dentro de la población.

Ahora bien, al generar estimaciones con base en una muestra de la población surge el **error muestral o de estimación**, el cual puede medirse a través de estadísticas como la varianza del estimador, el coeficiente de variación estimado (*cve*) el margen de error, entre otras. En el Icfes, el indicador de precisión más utilizado es el *cve*, el cual está dado por la siguiente expresión (Gutiérrez, 2009):

$$cve(\hat{T}) = \frac{\sqrt{\widehat{var}(\hat{T})}}{\hat{T}} * 100$$

Donde  $\hat{T}$  es la estimación del parámetro  $T$  y  $\widehat{var}(\hat{T})$  es el estimador de la varianza basado en la muestra seleccionada. Un valor muy alto en un *cve* indica que la estimación pierde confiabilidad debido a la variación de la muestra respecto a la estimación del parámetro. Se considera que un *cve* menor al 10% indica que la estimación es aceptable y es por tanto el umbral utilizado en el Icfes.

Cada uno de los subprocesos asociados a la selección de una muestra, puede ser una fuente de sesgo o de error. En la Figura 4, se listan algunas fuentes de error y de sesgo en los cuales se puede incurrir al seleccionar una muestra<sup>6</sup> (Särndal, 2003).

Es importante mencionar que la estimación del error mediante la formulación matemática es usualmente desconocida en diseños muestrales complejos. Sin embargo, en la actualidad se suelen aplicar técnicas computacionales o de re-muestreo<sup>7</sup> que permiten estimar estos valores. Por ejemplo, en pruebas como las Saber 3°, 5°, 7° y 9° se calcula este error a través de la metodología de réplicas balanceadas<sup>8</sup>.

6. Vale la pena mencionar que en un censo también se puede incurrir en errores diferentes al de muestreo.
7. El re-muestreo es una técnica que implica tomar muestras aleatorias repetidas de los datos disponibles y utilizarlas para estimar estadísticas y evaluar la precisión de los resultados. (Losilla, 1994)
8. Para profundizar sobre la metodología de réplicas balanceadas, puede consultar el siguiente enlace: [https://www.oecd-ilibrary.org/education/pisa-data-analysis-manual-sas-second-edition\\_9789264056251-en](https://www.oecd-ilibrary.org/education/pisa-data-analysis-manual-sas-second-edition_9789264056251-en)

FIGURA 4. Fuentes de error y de sesgo asociados a la selección de una muestra



## ¿Qué diseños de muestreo ha implementado el Icfes en las pruebas a gran escala?



Para la aplicación de las pruebas a gran escala se han implementado diseños muestrales estratificados multi-étapicos, en su mayoría. Por ejemplo, las pruebas Saber 3°, 5°, 7° y 9° fueron diseñadas para recolectar información sobre las habilidades en el ámbito cognitivo con pruebas como Matemáticas, Pensamiento Ciudadano, Ciencias Naturales y Educación Ambiental, y Competencias Comunicativas en Lenguaje: Lectura y Escritura.

Esta prueba también recolecta información sobre las condiciones económicas, sociales y emocionales de los estudiantes. A partir de esta información, se busca hallar evidencia sobre los factores que influyen en el aprendizaje de los estudiantes, esto con el fin de medir el cumplimiento de los objetivos educativos del país y adicional, implementar políticas públicas que aporten en el mejoramiento continuo de la educación.

Para la selección de esta muestra se realizó un muestreo multietápico que consta de tres etapas. Inicialmente se realiza una estratificación de todos los establecimientos dados por las siguientes variables:

- Zona en la que están ubicadas la mayoría de las sedes de un establecimiento: urbana y rural.
- Sector al que pertenece la institución: privado o público.
- Departamento.
- Oferta del establecimiento educativo: Oferta completa, que se refiere a los establecimientos que ofrecen los grados 3°, 5°, 7° y 9°; y oferta parcial, que se refiere a los establecimientos que ofrecen los grados 3° y 5°.

Para ejemplificar la construcción de los estratos para las pruebas Saber 3°, 5°, 7° y 9°, hay que tener en cuenta que cada estrato será una de todas las posibles combinaciones de las variables antes expuestas. Por ejemplo, para un departamento como Cundinamarca debemos combinar las demás variables de estratificación como lo son la zona, el sector y la oferta del establecimiento, resultando cada posible estrato para este departamento, como se puede ver en la Tabla 2.

**TABLA 2. Ejemplo de estratificación para las pruebas Saber 3°, 5°, 7° y 9° en Cundinamarca**

Departamento	Oferta	Sector	Zona
Cundinamarca	3°, 5°, 7° y 9°	Público	Urbano
	3° y 5°	Público	Urbano
	3°, 5°, 7° y 9°	Privado	Urbano
	3° y 5°	Privado	Urbano
	3°, 5°, 7° y 9°	Público	Rural
	3° y 5°	Público	Rural
	3°, 5°, 7° y 9°	Privado <sup>9</sup>	-
	3° y 5°	Privado	-

Este proceso se repite para cada uno de los departamentos del territorio nacional y de esta manera se obtienen todos los estratos considerados en las pruebas Saber 3°, 5°, 7° y 9°. Hay que tener en cuenta que la estratificación no es una etapa de muestreo pues no

<sup>9</sup> Es preciso señalar que no hay desagregación Rural-Privado dado que, por las dinámicas propias del país, hay muy pocos establecimientos que cumplen esta condición

se hace ningún tipo de selección, solo se subdivide la población según sus características comunes.

Después de realizar la estratificación, se procede a realizar el muestreo en tres etapas que se desarrolla de la siguiente manera:

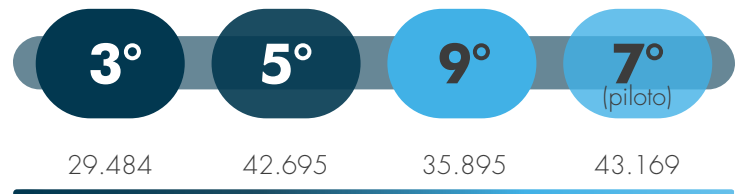
- **Primera etapa:** Se hace una selección proporcional al tamaño del establecimiento educativo dado por el número de estudiantes matriculados en los grados 3°, 5°, 7° y 9°.
- **Segunda etapa:** En cada uno de los establecimientos seleccionados en la primera etapa, se eligen sedes bajo un muestreo aleatorio simple, separando las sedes por oferta completa o incompleta.
- **Tercera etapa:** Todos los estudiantes matriculados en los grados de interés antes mencionados y que pertenecen a las sedes seleccionadas en la segunda etapa de muestreo, presentan las pruebas de Matemáticas y Lectura. Sin embargo, para 5°, 7° y 9° se realiza una asignación sistemática de cuadernillos donde la mitad de los estudiantes presentan las pruebas de Acciones y Actitudes, y Pensamiento Ciudadano; mientras que la otra mitad presentan las pruebas de Ciencias Naturales y Escritura.

Después de realizar la selección de la muestra, se aplica la prueba a los estudiantes seleccionados. De esta forma, para la aplicación del 2021, se obtuvo una muestra efectiva de 108.074 estudiantes en los grados 3°, 5° y 9° que se encontraban distribuidos en 598 establecimientos y a su vez en 1.230 sedes. Con los resultados de la prueba, se obtuvieron estimaciones de puntajes promedios que fueron insesgadas y precisas, lo cual garantizó la representatividad de los resultados a nivel nacional y regional. Adicionalmente, se garantizaron estimaciones representativas a nivel de zona, sector y departamento.

Para la aplicación de la prueba en 2021 no se obtuvo estimaciones del grado 7°, pues en este año se implementó la prueba por primera vez para este grado.

Por esto, en esta aplicación se realizó el pilotaje para 7°, con el fin de robustecer el banco de preguntas que determinará la estructura de la prueba para las siguientes aplicaciones.

**FIGURA 5.** Muestra seleccionada para las pruebas Saber 3°, 5°, 7° y 9°, por cada grado.


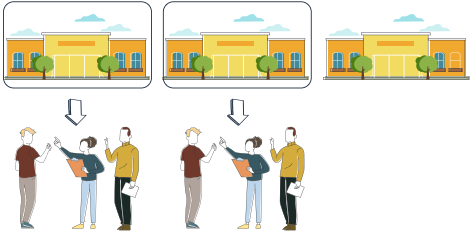

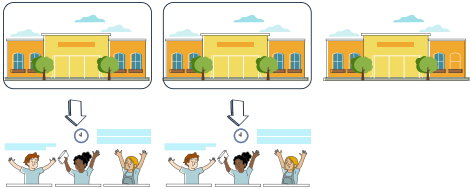


Así como las pruebas Saber 3°, 5°, 7° y 9° tienen un objetivo particular, cada una de las pruebas a gran escala en las que participa el Icfes también tienen su propio objetivo. A continuación, en la Tabla 3, se encuentran las generalidades de cada una de las pruebas a gran escala, su objetivo, la población objetivo, el diseño muestral aplicado y las etapas de muestreo con un resumen del diseño aplicado en cada etapa.

TABLA 3. Pruebas a gran escala en las que ha participado o desarrollado el Icfes, que han contado con diseños muestrales probabilísticos

Prueba	Objetivo	Población objetivo	Diseño muestral	Etapas <sup>10</sup>
	<p>Las pruebas Saber 3°, 5°, 7° y 9°, tiene como objetivo brindar información al país sobre el desempeño de los y las estudiantes en áreas fundamentales de educación básica primaria y secundaria, así como proporcionar datos sobre las habilidades socioemocionales y algunos factores que pueden asociarse al desarrollo de los aprendizajes.</p>	<p>Estudiantes de 3.º, 5.º, 7.º y 9.º de calendario A</p>	<p>Diseño probabilístico estratificado en tres etapas.</p>	<p><b>Primera etapa:</b> Selección de establecimientos educativos con probabilidades de inclusión proporcionales al número de estudiantes.</p> <p><b>Segunda etapa:</b> Muestreo aleatorio simple de sedes educativas.</p> <p><b>Tercera etapa:</b> Censo de estudiantes y selección sistemática para asignar cuadernillos.</p>
	<p>El Programa Internacional para la Evaluación de los Estudiantes, PISA (por sus siglas en inglés), tiene como objetivo determinar en qué medida los estudiantes de 15 años, independientemente del grado que estén cursando, han adquirido los conocimientos y competencias esenciales para afrontar los retos de la vida adulta.</p>	<p>Estudiantes de 15 años, que estén cursando entre 7º a 11º.</p>	<p>Diseño probabilístico estratificado en dos etapas (OECD 2018a)</p>	<p><b>Primera etapa:</b> Selección de sedes educativas con probabilidades de inclusión proporcionales al número de estudiantes. Cuando una ciudad desea representatividad, se añade un estrato para estos lugares, específicamente.</p> <p><b>Segunda etapa:</b> Selección de estudiantes de forma aleatoria simple.</p>

10. Corresponden a las etapas de selección dentro de cada estrato

Prueba	Objetivo	Población objetivo	Diseño muestral	Etapas <sup>10</sup>
	<p>El Estudio Internacional de Enseñanza y Aprendizaje (TALIS, por sus siglas en inglés), promovido por la OCDE, tiene como objetivo recolectar información sobre el aprendizaje y las condiciones de trabajo de los profesores en los colegios.</p>	<p>Docentes y trabajadores del sector educativo</p>	<p>Diseño probabilístico estratificado de dos etapas (Ainley y Carstens, 2018)</p>	 <p>Primera etapa: Selección aleatoria simple de sedes.</p> <p>Segunda etapa: Selección aleatoria simple de docentes.</p>
	<p>El Estudio Regional Comparativo y Explicativo (ERCE), tiene como objetivo obtener datos comparativos sobre el desempeño escolar en América Latina y el Caribe, específicamente en las áreas de ciencias naturales, matemáticas y lectura, así como analizar los factores asociados al desempeño escolar para retroalimentar la política educativa. A la fecha se han realizado 4 estudios regionales: PERCE, SERCE, TERCE y ERCE.</p>	<p>Estudiantes de 3.º y 6.º. Adicionalmente, los rectores y cuidadores responden una encuesta</p>	<p>Diseño probabilístico estratificados de dos etapas (LLECE, 2021)</p>	 <p>Primera etapa: Selección de sedes educativas con probabilidades de inclusión proporcionales al número de estudiantes. Cuando una ciudad desea representatividad, se añade un estrato para estos lugares, específicamente.</p> <p>Segunda etapa: Selección de salones de forma aleatoria simple.</p>

Prueba	Objetivo	Población objetivo	Diseño muestral	Etapas <sup>10</sup>
	<p>El Estudio Internacional sobre Habilidades Sociales y Emocionales (SSES, por sus siglas en inglés) busca evaluar las habilidades sociales y emocionales en estudiantes por medio de auto-reportes y recogiendo información de su contexto. (OECD, 2021)</p>	<p>Estudiantes de 10 años que estén en 2° o superior.</p> <p>Estudiantes de 15 años que estén en 7° o superior.</p> <p>Adicionalmente, se aplican cuestionarios a padres, rectores y profesores.</p>	<p>Diseño probabilístico en dos etapas (OECD, 2021)</p>	<p><b>Primera etapa:</b> Selección de sedes educativas con probabilidades de inclusión proporcionales al número de estudiantes.</p> <p><b>Segunda etapa:</b> Selección de estudiantes de forma aleatoria simple.</p>
	<p>El Estudio Internacional sobre Educación Cívica y Ciudadana, (ICCS, por sus siglas en inglés), busca evaluar las formas cómo los jóvenes están preparados para desarrollar su rol de ciudadanos en cuanto a sus conocimientos, competencias y actitudes.</p>	<p>Jóvenes de 14 años, que estén cursando 8.º</p>	<p>Diseño probabilístico estratificado en dos etapas (IEA, 2018)</p>	<p><b>Primera etapa:</b> Selección de sedes educativas con probabilidades de inclusión proporcionales al tamaño (número de estudiantes).</p> <p><b>Segunda etapa:</b> Selección de salones con muestreo aleatorio sistemático para los estudiantes y muestreo aleatorio simple para los docentes.</p>
	<p>El Estudio Internacional de Progreso en Comprensión Lectora, PIRLS, por sus siglas en inglés, evalúa el rendimiento en lectura. Se inauguró en el 2001 como seguimiento del estudio de alfabetización de la lectura de 1991 de la IEA.</p>	<p>Estudiantes que estén cursando 4.º escolaridad</p>	<p>Diseño probabilístico estratificado de dos etapas (LaRoche et al., 2017)</p>	<p><b>Primera etapa:</b> Selección de sedes educativas con probabilidades de inclusión proporcionales al número de estudiantes.</p> <p><b>Segunda etapa:</b> Selección de salones de forma aleatoria simple.</p>



Prueba	Objetivo	Población objetivo	Diseño muestral	Etapas <sup>10</sup>
	<p>El Estudio Internacional de Tendencias en Matemáticas y Ciencias, TIMSS, por sus siglas en inglés evalúa a estudiantes de cuarto y octavo grado y tiene como propósito generar información para mejorar los procesos de enseñanza y aprendizaje de las matemáticas y las ciencias.</p>	<p>Estudiantes que estén cursando 4.º y 8.º de escolaridad</p>	<p>Diseño probabilístico estratificado de dos etapas (LaRoche et al., 2017)</p>	<div style="display: flex; align-items: center;">  <div style="margin-left: 20px;"> <p><b>Primera etapa:</b> Selección de sedes educativas con probabilidades de inclusión proporcionales al número de estudiantes.</p> <p><b>Segunda etapa:</b> Selección de estudiantes de forma aleatoria simple.</p> </div> </div>

## ¿Qué recomendaciones se deben tener en cuenta para el procesamiento de los datos de las pruebas a gran escala?



Es muy importante tener en cuenta que, para el uso de los datos de cualquiera de las pruebas nacionales o internacionales a gran escala en las que participa el Icfes, es totalmente necesario hacer uso de cada una de las variables asociadas al diseño muestral y el modelo de calificación, empezando con los pesos de muestreo que permiten expandir la información de la muestra para poder hablar de los resultados a nivel de las desagregaciones contempladas en el diseño de muestreo. Adicionalmente, según la particularidad de cada prueba y el modelo de calificación, algunas de estas hacen uso de valores plausibles, los cuales son múltiples imputaciones de la habilidad de cada estudiante.

Además, dada la complejidad de algunos de los diseños muestrales presentados previamente, se generan réplicas balanceadas para calcular los errores de muestreo descritos anteriormente. Para usar todas estas variables, existen algoritmos que se deben aplicar bajo unas reglas de procesamiento. Por lo tanto, cualquier estimación que sea generada por fuera de estas reglas y parámetros carece de representatividad y generalización. Para hacer uso correcto de los datos, existe una guía de procesamiento que permite orientar este proceso. Los invitamos a consultarla en el siguiente enlace:

[MANUAL PISA](#)

## Conclusiones

El Icfes ha desarrollado y participado en diversas pruebas diagnósticas de educación conocidas como pruebas a gran escala. Estas evaluaciones han sido aplicadas mediante técnicas de muestreo basadas en estándares internacionales, que, gracias a su adecuada implementación, permiten obtener evidencias para generar políticas públicas pertinentes y oportunas, orientadas a mejorar la calidad educativa a nivel nacional.

Existen diferentes tipos de diseños probabilísticos que pueden ser aplicados según el objetivo de la investigación y las características de la población en términos estructurales y de variabilidad. En las pruebas que desarrolla y en las que participa el Instituto, se suele utilizar el diseño muestral en varias etapas teniendo en cuenta la estructura del sistema educativo del país.

Una correcta aplicación de un diseño de muestreo probabilístico puede proporcionar una representación precisa y confiable de la población que se desea analizar,

lo que a su vez permite realizar inferencias robustas acerca del fenómeno estudiado. Al seleccionar adecuadamente las muestras y controlar factores como el tamaño de la muestra y el uso de instrumentos de medición adecuados, se puede llegar a conclusiones generalizables y respaldadas por datos sólidos, y así garantizar que los resultados obtenidos de una muestra sean precisos y que sean representativos de toda la población.

Por último, al hacer una medición de una porción de la población en lugar de examinar su totalidad, es decir, al seleccionar una muestra, es importante reconocer que, de manera inmediata, se incurre en un error de muestreo, así como en errores originados por otras fuentes. Sin embargo, estos errores no deben subestimarse, ya que su magnitud y efecto pueden ser evaluados a través de medidas de calidad como el coeficiente de variación u otras herramientas estadísticas adecuadas que brindan una idea de la precisión de los resultados obtenidos.

---

## Referencias

*Ainley, J. y Carstens, R. (2018).* Teaching and Learning International Survey (TALIS) 2018 Conceptual Framework. *OECD Education Working Papers*, (187).

[Link al documento](#)

*Bhandari, P. (2020, 27 de noviembre).* Parameter vs Statistic I Definitions, Differences & Examples. Scribbr.

[Link al documento](#)

*Australian Bureau (2023a/s.f.)* Census and Sample

[Link al documento](#)

*Australian Bureau (2023b/s.f.)* Sample Design

[Link al documento](#)

*Departamento Administrativo Nacional de Estadística [DANE]. (2005).* Metodología de diseño muestral. Encuesta desempeño Institucional (EDI). Dirección de Regulación, Planeación, Estandarización y Normalización.

[Link al documento](#)

*Gutiérrez, A. (2009).* Estrategias de muestreo, diseño de encuestas y estimación de parámetros (2.ª ed.). Ediciones de la U.

*International Association for the Evaluation of Educational Achievement [IEA]. (2018).* International Civic and Citizenship Education Study 2016. Technical Report.

[Link al documento](#)

*LaRoche, S., Joncas, M. & Foy, P. (2017).* Chapter 3: Sample design in PIRLS 2016. En M. Martin, I. Mullis y M. Hooper (eds.), *Methods and procedures in PIRLS 2016* (pp. 3.1-3.34). Boston College, TIMSS & PIRLS International Study Center [Link al documento](#)

*Laboratorio Latinoamericano de Evaluación de la Calidad de la Educación [LLECE]. (2021).* Estudio regional comparativo y explicativo (ERCE 2019). Reporte de resultados Colombia. Unesco.

[Link al documento](#)

*Losilla, J. (1994).* MonteCarlo: toolbox de Matlab. Herramientas para un laboratorio de estadística fundamentado en técnicas Monte Carlo (tesis de doctorado, Universitat Autònoma de Barcelona). Dipòsit Digital de Documents de la UAB.

[Link al documento](#)

*Organisation for Economic Co-operation and Development [OECD]. (2018)* Chapter 4. Sample Design. Target Population and Overview of the Sampling Design. En PISA 2018 *Technical Report*. OECD Publishing

[Link al documento](#)

*Organisation for Economic Co-operation and Development [OECD]. (2021).* OECD Survey on Social and Emotional Skills Technical Report. OECD.

[Link al documento](#)

*Särndal, C.-E., Swensson, B., y Wretman, J. (2003).* Model assisted survey sampling. Springer Science & Business Media.

# SABER AL > DETALLE

**¿CÓMO SE REALIZAN  
LOS MUESTREOS  
PROBABILÍSTICOS DE LAS  
PRUEBAS A GRAN ESCALA  
QUE DESARROLLA Y EN  
LAS QUE PARTICIPA EL  
ICFES?**



**Educación**



@icfescol



ICFES



icfescol



YouTube: ICFES

Instituto Colombiano para la Evaluación de la Educación, ICFES  
Oficinas Calle 26 No. 69 - 76. Torre 2, pisos 15 - 18  
Edificio Elemento, Bogotá. Colombia