

Plan Nacional para la Evaluación de los Aprendizajes (PLANEA)





Plan Nacional para la Evaluación de los Aprendizajes (PLANEA)





Plan Nacional para la Evaluación de los Aprendizajes (PLANEA)
es una publicación digital del Instituto Nacional
para la Evaluación de la Educación.

Su elaboración estuvo a cargo de la Unidad de Evaluación
del Sistema Educativo Nacional.

Agosto, 2015

Contenido

El Plan Nacional para la Evaluación de los Aprendizajes (PLANEA)	6
I. Antecedentes	7
Fundamentos legales	7
La evaluación del aprendizaje a gran escala	7
Las pruebas de Estándares Nacionales, EXCALE y ENLACE	9
Estudio de validación de ENLACE básica, ENLACE media superior y EXCALE	10
II. Propósitos de PLANEA	11
Elementos que orientan la definición de PLANEA	12
III. Características generales de PLANEA	13
Modalidades de evaluación y esquema de aplicación	14
Población objetivo	15
Evaluación de aprendizajes clave	15
Evaluación de habilidades socioafectivas	16
La contextualización de las evaluaciones	17
Participación de diferentes grupos de expertos en el diseño y elaboración de los instrumentos de evaluación	18
IV. Precisiones sobre las modalidades de PLANEA	19
Evaluación de logro referida al Sistema Educativo Nacional (ELSEN)	20
Evaluación del logro referida a los centros escolares (ELCE)	21
Evaluación diagnóstica censal (EDC)	25
Participación de diferentes actores en las modalidades de PLANEA	26
V. Difusión y uso de resultados	27
Reportes de la evaluación del logro referida a los centros escolares	29
VI. Revisión y seguimiento del Plan	30

Anexo 1. Procedimiento previsto de aplicación	31
Anexo 2. Descripción de las escalas de habilidades para la convivencia escolar	34
Anexo 3. Estructura de los cuestionarios de contexto	37
Anexo 4. Aspectos centrales del diseño de la muestra para las Evaluaciones referidas al Sistema Educativo Nacional (ELSEN)	42
Anexo 5. Síntesis del análisis de la calidad de los resultados de las Evaluaciones del logro referidas a los centros escolares (ELCE) de PLANEA	53

El Plan Nacional para la Evaluación de los Aprendizajes (PLANEA)

La evaluación de los aprendizajes —tanto la que se lleva a cabo al interior de las escuelas como la que se aplica de manera externa— es una herramienta útil para la mejora educativa. Las evaluaciones internas, cuando son realizadas por los docentes con un enfoque formativo, son de gran relevancia porque permiten identificar lo que aprende cada estudiante, además de ser un medio para obtener información continua sobre qué tan bien se organizaron las condiciones para el aprendizaje y qué tan buenos fueron los procedimientos instruccionales utilizados. Lo anterior se debe a que los docentes pueden emplear una gran diversidad de métodos de evaluación a lo largo del tiempo, en condiciones cotidianas de aula, y con un conocimiento pormenorizado de las circunstancias en que se llevan a cabo los aprendizajes de sus alumnos.

Las evaluaciones externas complementan a las internas y ofrecen un panorama amplio sobre lo que han aprendido los niños de diferentes entornos sociales y escolares. En ellas típicamente participan grandes cantidades de alumnos (por lo que se denominan evaluaciones a gran escala), suelen aplicarse al final de un proceso educativo (por ejemplo, al término de un ciclo escolar), consideran los aprendizajes que son comunes a todos los evaluados (tales como los contenidos fundamentales del plan de estudios) y son estandarizadas para permitir la comparación.

Con el fin de conocer qué tanto los estudiantes mexicanos dominan aprendizajes clave, el Instituto Nacional para la Evaluación de la Educación (INEE) ha desarrollado el Plan Nacional para la Evaluación de los Aprendizajes (PLANEA) a partir del cual se obtiene información complementaria para apoyar la mejora educativa. En la elaboración de este Plan se ha considerado como antecedente el análisis de las fortalezas y limitaciones de las pruebas EXCALE y ENLACE y se atiende a las recomendaciones derivadas de un estudio que se llevó a cabo a petición del INEE durante 2013 y 2014 por un grupo de especialistas de diferentes instituciones.¹

¹ Martínez Rizo, F. (Coord.) (2015). *Las pruebas ENLACE y EXCALE. Un estudio de validación*. México: INEE.

El presente documento describe las características principales de PLANEA. Se organiza en seis apartados: en el primero se exponen los fundamentos legales y los antecedentes de evaluación de PLANEA. En el segundo se describen los elementos que orientan la definición de PLANEA y sus propósitos. En el tercer apartado se explican las características más relevantes de las modalidades de evaluación que componen el Plan. El cuarto apartado contiene información más detallada sobre las características de cada modalidad. El quinto se refiere a la difusión y el uso de los resultados de las evaluaciones. Finalmente, en el sexto apartado se presentan elementos para la estrategia de revisión y seguimiento del Plan.

I. Antecedentes

Fundamentos legales

Las atribuciones del INEE, establecidas en el artículo 3° de la Constitución Política de los Estados Unidos Mexicanos, el artículo 29 de la Ley General de Educación, y el artículo 27 de la Ley del INEE, relacionadas con el desarrollo de estas evaluaciones son:

1. Diseñar y realizar mediciones de los componentes del Sistema Educativo Nacional (SEN), entre otros, el aprendizaje de los alumnos.
2. Expedir los lineamientos a los que se sujetarán las autoridades educativas para realizar las funciones de evaluación que les correspondan.
3. Convenir con las autoridades educativas la aplicación de instrumentos destinados a evaluar de manera confiable, válida y periódica el nivel de aprendizaje alcanzado por los alumnos.

La evaluación del aprendizaje a gran escala

En las primeras décadas del siglo XX, con el desarrollo de los tests para la evaluación individual de capacidades humanas y para la selección o clasificación de estudiantes en las escuelas, se inició una era de evaluaciones estandarizadas con diferentes propósitos aplicadas a grandes poblaciones. En la década posterior a la segunda guerra mundial, las evaluaciones grupales estandarizadas empezaron a aplicarse en las escuelas con propósitos de evaluación curricular y en mucho menor grado para monitorear los resultados de los sistemas educativos.²

El desarrollo y aplicación de pruebas de gran escala para evaluar el logro escolar de niños y jóvenes en los sistemas educativos data de principios de los años sesenta del siglo veinte. Los primeros estudios sobre el rendimiento en matemáticas y ciencias, comparando los resultados de varios países, fueron coordinados por la Asociación Internacional para la

² Goslin, Epstein & Hallock, 1965.

Evaluación de Logro Educativo (IEA, por sus siglas en inglés) en 1964. Esta organización fue promovida inicialmente por un grupo de investigadores del Instituto de Educación de la UNESCO en Hamburgo, Alemania, para realizar estudios comparativos sobre el logro educativo en diferentes países. Después, la IEA se formalizó como una entidad internacional en 1967 en Bruselas, Bélgica. Desde entonces ha desarrollado diversos programas de pruebas estandarizadas como *Trends in International Mathematics and Science Study* (TIMSS), *Progress in International Reading Literacy Study* (PIRLS), *International Civic and Citizenship Study* (ICCS), *International Computer and Information Literacy Study* (ICILS), y los estudios del Laboratorio Latinoamericano de Evaluación de la Calidad de la Educación (LLECE), que se aplican en un número creciente de países.

Un referente para la aplicación de pruebas estandarizadas a gran escala es el Programa Nacional de Evaluación del Avance Educativo (NAEP, por sus siglas en inglés), que inició en Estados Unidos en 1969. Este programa que ha tenido continuidad durante las últimas cuatro décadas —con un auge importante en los años ochenta y noventa— evalúa el logro del aprendizaje en muestras de alumnos cada tres o cuatro años, para dar cuenta de las tendencias de este indicador en el sistema educativo norteamericano a lo largo del tiempo.

En las últimas décadas se han reconsiderado diversas aproximaciones y se han analizado prácticas de evaluación educativa a gran escala, a la luz de los avances en las ciencias cognitivas, medición educativa y métodos de análisis estadístico, para mejorar su diseño, implementación y el uso válido de los resultados.³

Nuestro país tomó parte en la aplicación de pruebas estandarizadas a nivel nacional en matemáticas y ciencias en el programa TIMSS en 1995, aunque posteriormente se retiró del estudio.

En el año 2000 México participó por primera vez en el proyecto PISA de la Organización para la Cooperación y Desarrollo Económico (OCDE) y subsecuentemente en 2003, 2006, 2009, 2012 y 2015. Este programa internacional ha permitido fomentar la aplicación sistemática de evaluaciones de logro educativo, en particular para conocer el aprendizaje que logran grupos de estudiantes de 15 años de cada país participante, en lectura, matemáticas y ciencias.

Otras experiencias en el diseño y aplicación de pruebas estandarizadas para evaluar a grandes grupos de alumnos en México, se han realizado desde hace varias décadas en diversas instituciones universitarias; el Centro Nacional de Evaluación para la Educación Superior (CENEVAL), desde 1994, las lleva a cabo en los niveles de educación media superior y superior, con propósitos de admisión y de diagnóstico académico, entre otros objetivos.

³ Pellegrino, Chudowsky y Glaser, 2001; Mislevy, 2006; Simon, Ercikan, y Rousseau, 2013.

Las pruebas de Estándares Nacionales, EXCALE y ENLACE

En México, el uso de evaluaciones de logro con alcance nacional en educación básica es relativamente reciente. Al respecto pueden destacarse tres grandes proyectos evaluativos: 1) el desarrollo de las pruebas de Estándares Nacionales que desarrolló la Secretaría de Educación Pública (SEP) a finales de los años noventa como un primer esfuerzo a nivel nacional; 2) las evaluaciones de aprendizaje de gran escala diseñadas por el INEE, conocidas como Exámenes para la Calidad y Logro Educativos (EXCALE), que iniciaron sus aplicaciones en 2005; y 3) las Evaluaciones Nacionales de Logro Académico en Centros Escolares (ENLACE), desarrolladas por la SEP a partir de 2006.

Las evaluaciones de Estándares Nacionales fueron diseñadas con el propósito de contar con mediciones que permitieran conocer el estado del sistema educativo y su evolución en el tiempo, pero sus resultados se presentaban con altos niveles de agregación, es decir, con referencia a grupos muy amplios de estudiantes y a dimensiones evaluadas muy generales. Esas pruebas no se diseñaron con el propósito de evaluar los contenidos curriculares, por lo que no proporcionaban información acerca del dominio de los contenidos de los planes de estudio que lograban los alumnos; la información que se suministraba a las autoridades educativas era limitada debido a que no daba cuenta del logro en relación con los propósitos del currículum. En síntesis, las evaluaciones de Estándares Nacionales: a) no permitían construir una visión general del aprendizaje de los estudiantes como resultado de su escolarización formal, b) no daban información acerca de los puntos fuertes y débiles del aprendizaje de los estudiantes en las asignaturas más importantes y, c) no permitían las comparaciones del rendimiento escolar, ni la posibilidad de observar las tendencias de los aprendizajes a lo largo del tiempo.⁴

Por otro lado, los EXCALE eran pruebas criterioles y con un diseño matricial, y estaban alineadas al currículum nacional, lo que permitía evaluar un dominio amplio de contenidos de los programas de estudio evaluados; su escalamiento se basó en la Teoría de Respuesta al Ítem y en criterios de la Teoría Clásica de los Tests para cuidar la calidad de las pruebas aplicadas a muestras a nivel nacional y estatal cada cuatro años en tercero de preescolar, tercero y sexto de primaria y tercero de secundaria.⁵ Las evaluaciones EXCALE se aplicaron de manera ininterrumpida durante los últimos diez años.

Por su parte, la prueba ENLACE fue diseñada para aplicarse de manera censal y tuvo como propósito original informar a padres de familia y a docentes sobre el rendimiento de los estudiantes de tercero de primaria a tercero de secundaria y del último grado de media superior en Español, Matemáticas y una tercera asignatura seleccionada cada año (por ejemplo Ciencias). Sin embargo, esa intención original se transformó en otra de rendición de cuentas de escuelas y docentes, a la que se asociaron propósitos no previstos y altas consecuencias como la publicación de *rankings* escolares, el otorgamiento de estímulos económicos a los docentes y la premiación a los alumnos con mejores puntuaciones en este tipo de exámenes.

⁴ Backhoff y Martínez-Rizo, 2004.

⁵ Backhoff, Sánchez, Peón, Monroy, 2006.

Estudio de validación de ENLACE básica, ENLACE media superior y EXCALE

En 2013, el INEE solicitó a un comité de expertos un estudio para analizar la validez de las pruebas ENLACE y EXCALE. Además, organizó el Seminario Internacional “Hacia una nueva generación de evaluaciones estandarizadas”, a fin de fundamentar el desarrollo de las nuevas evaluaciones nacionales del aprendizaje, tomando en consideración las experiencias exitosas de otros países.

Los estudios de validación de las pruebas ENLACE y EXCALE⁶ permiten reconocer los aciertos y debilidades de las dos pruebas más destacadas en la experiencia nacional de evaluación estandarizada de los últimos diez años. Las conclusiones de este estudio apuntan a que las pruebas EXCALE tuvieron la virtud de haber evaluado periódicamente el aprendizaje de los estudiantes de manera válida, confiable y estable, pero sus resultados tuvieron poca visibilidad social e insuficiente presencia en la toma de decisiones de política educativa.

Por su parte, la fortaleza de ENLACE radicó en haber proporcionado información sobre los resultados de aprendizaje a nivel de alumno y de escuela, y en haber logrado una amplia visibilidad entre las autoridades educativas y diversos grupos de la sociedad civil; sin embargo, tanto la aplicación excesiva de evaluaciones (muchos grados, todos los años y censales) como las consecuencias asociadas con los resultados para escuelas y docentes (*rankings* de escuelas e incentivos económicos a los profesores) provocaron que el indicador de aprendizaje se corrompiera, los resultados fueran alterados y se desgastara la incipiente confianza y credibilidad en las evaluaciones por parte de distintos actores sociales y educativos.

Las experiencias nacionales de evaluación de logro educativo y la participación de México en programas internacionales con objetivos similares han proporcionado lecciones muy importantes. Para que las evaluaciones nacionales del aprendizaje cumplan con su propósito es necesario: a) que se cuide el rigor metodológico en su diseño, construcción y aplicación; b) que se norme el uso e impacto de los resultados y se garantice que la información que ofrecen sus resultados se corresponda con los propósitos para los que fueron diseñadas; y, c) que se garantice la precisión y la comparabilidad de las medidas, a fin de lograr que la información que se dé sobre los posibles cambios a lo largo del tiempo sea lo más confiable posible.

⁶ Martínez Rizo, F. (Coord.) Op. Cit.

II. Propósitos de PLANEA

PLANEA tiene como propósito general conocer la medida en que los estudiantes logran el dominio de un conjunto de aprendizajes esenciales en diferentes momentos de la educación obligatoria.

Los resultados de las evaluaciones de PLANEA servirán para la mejora educativa, a partir de:

- Informar a la sociedad sobre el estado que guarda la educación en términos del logro de aprendizaje de los estudiantes y de la equidad (o inequidad) que existe en los resultados educativos.
- Aportar a las autoridades educativas información relevante para el monitoreo, la planeación, programación y operación del sistema educativo y sus centros escolares.
- Ofrecer información pertinente, oportuna y contextualizada a las escuelas y a los docentes, que ayude a mejorar sus prácticas de enseñanza y el aprendizaje de todos sus estudiantes.
- Contribuir al desarrollo de directrices para la mejora educativa con información relevante sobre los resultados educativos y los contextos en que se dan.

Los resultados de las evaluaciones PLANEA no deberán utilizarse para juzgar el desempeño de los docentes, realizar *rankings* de escuelas, justificar procesos punitivos u otros de control administrativo sobre estudiantes, docentes o escuelas, por cuatro razones fundamentales:

1. No son sus propósitos previstos, ni es de interés del Plan servir a estos fines. Además, estos propósitos podrían, por su naturaleza, contraponerse con el de ofrecer información válida y confiable para contribuir a la mejora, disminuyendo la capacidad de PLANEA para cumplir con este último.
2. El diseño de las tres modalidades que componen el Plan no permite obtener información suficiente para ninguno de estos usos. Para juzgar el desempeño de los docentes deben incorporarse diversos elementos basados en varias fuentes, como por ejemplo, instrumentos para observación en aula aplicados por personas capacitadas para ello, cuestionarios de percepción de la clase aplicados a los alumnos y ganancia en el logro de los estudiantes,⁷ entre otros.
3. La información que se ofrece a nivel de centro escolar no es suficiente para hacer un ordenamiento justo de las escuelas, puesto que los resultados de las evaluaciones tienen múltiples causas que actúan simultáneamente y que deberían considerarse al analizar la eficacia que se alcanza debido a la acción de los integrantes de un centro escolar. Este propósito y las herramientas metodológicas que se requerirían, exceden las posibilidades de PLANEA.

⁷ Cfr. Bill & Melinda Gates Foundation (2013). *Ensuring Fair and Reliable Measures of Effective Teaching. Culminating Findings from the MET Project's Three-Year Study.*

4. Algunos de estos usos no parecen especialmente útiles para la mejora de la educación. Por ejemplo, se dice que hacer *rankeos* de escuelas las coloca en una posición de exigencia y que esto ayuda a mejorar la educación, pero en numerosos contextos,⁸ incluyendo el nacional, se ha mostrado que más bien se genera el fenómeno conocido como inflación de resultados, es decir, el crecimiento desmedido y artificial de las puntuaciones de una prueba, como reacción a las consecuencias asociadas a los resultados más que a cambios reales.

Elementos que orientan la definición de PLANEA

La revisión de las experiencias de evaluación anteriores, particularmente ENLACE y EXCALE, así como la búsqueda de diseños que cumplan con los propósitos expresados y eviten los usos inadecuados enunciados, condujo a considerar los siguientes elementos para la definición de PLANEA:

1. Se toman en cuenta los aciertos y errores de las experiencias nacionales e internacionales en materia de evaluaciones estandarizadas para evaluar sistemas educativos.
2. Se centra en la rendición de cuentas a nivel nacional (qué tanto aprenden los estudiantes respecto a lo que deberían aprender, y en qué medida hay inequidad entre los aprendizajes de distintos grupos de población), y en el uso pedagógico a nivel de escuela (cómo puede la comunidad escolar mejorar los aprendizajes de todos sus estudiantes, una vez reconocidos elementos del contexto que dificultan o facilitan el logro).
3. Se utiliza un esquema complementario de los diseños muestral y censal, que mide los mismos aprendizajes.
4. La unidad de análisis para las evaluaciones censales es la escuela, mientras que para las evaluaciones muestrales son los estratos escolares seleccionados (país, entidades, tipo de servicio, etcétera).
5. Se utiliza un esquema de coordinación INEE–SEP–Autoridades estatales–Escuelas.
6. Se incorporan procedimientos para identificar el sesgo de las evaluaciones que pudiera resultar de la diversidad de género, étnica, cultural y lingüística, y para evitar comparaciones injustas.
7. Los resultados de las evaluaciones de logro escolar no tienen consecuencias adversas sobre estudiantes, docentes ni escuelas. Tampoco se considera apropiado otorgar premios por el buen desempeño en las pruebas.

⁸ Holcombe, R., Jennings, J., & Koretz, D. (2013). The roots of score inflation: An examination of opportunities in two states tests. In G. Sunderman (Ed.), *Charting reform, achieving equity in a diverse nation*. Greenwich, CT: Information Age Publishing.

8. Se implementan mecanismos para evitar los efectos inflacionarios de los resultados de las evaluaciones.
9. Hay un estricto control en la aplicación de las evaluaciones, y se señalan los casos en que los resultados no son confiables.
10. Se busca un uso eficiente de los recursos que se destinarán a las tareas evaluativas, de forma conjunta con las autoridades educativas.

En la siguiente sección se presentan las características de PLANEA.

III. Características generales de PLANEA

Los propósitos de PLANEA suponen que diferentes tipos de información se pongan a disposición de una serie de actores educativos, para su uso. Existen tres modalidades en el diseño de los instrumentos y las aplicaciones de PLANEA, lo que obedece a que cada una de ellas resulta más adecuada para obtener cierto tipo de información y por lo tanto, para alcanzar mejor alguno de los propósitos ya enunciados. A continuación se presenta una breve descripción de las modalidades, y más adelante se caracterizarán de manera más completa.

1. **Evaluación del logro referida al SEN (ELSEN):** Tiene el propósito de informar a la sociedad sobre el estado que guarda la educación en términos del logro de aprendizaje de los estudiantes, aportar información relevante y utilizable a las autoridades educativas nacionales y estatales para el monitoreo, la planeación, programación y operación del sistema educativo, y contribuir con información relevante sobre los resultados educativos y sus contextos, a nivel de sistema, para el desarrollo de directrices para la mejora educativa. El INEE aplica los instrumentos de esta modalidad a muestras representativas de estudiantes a nivel nacional, estatal y de diferentes estratos escolares, de los grados terminales de la educación obligatoria, en ciclos de cuatro años. Los instrumentos de esta modalidad tienen un arreglo matricial (que se describe más adelante) que permite evaluar una cantidad amplia de aprendizajes clave.
2. **Evaluación del logro referida a los centros escolares (ELCE):** Su propósito es ofrecer información contextualizada para la mejora de los procesos de enseñanza en los centros escolares y aportar información para el monitoreo, la planeación, programación y operación del sistema educativo, y a nivel directivo de las escuelas. Evaluará una cantidad reducida de aprendizajes clave y sus instrumentos serán aplicados anualmente por la SEP en coordinación con las autoridades educativas estatales, en todas las escuelas del país, para los grados terminales de primaria, secundaria y educación media superior. Aunque la aplicación se hará en todas las escuelas, no será administrada a todos los alumnos de los grados correspondientes. Para la educación básica se harán muestras de alumnos por escuela en los planteles en que haya más de 35 alumnos del grado a evaluar (aproximadamente 28% de las escuelas del país).

3. Evaluación diagnóstica censal (EDC): Tiene el propósito de ofrecer información pertinente, oportuna y contextualizada a las escuelas, que ayude a los docentes a mejorar sus prácticas de enseñanza y el aprendizaje de sus alumnos, particularmente en lo relacionado con la planeación de inicio de cursos, basada en un diagnóstico de los aprendizajes de todos los alumnos. La SEP entregará pruebas estandarizadas a los docentes de cuarto grado de primaria al inicio del ciclo escolar, para que las apliquen, analicen los resultados a partir de una guía y los utilicen como elementos para la planificación del trabajo en el aula. Estas pruebas, bajo el control de las escuelas y los profesores, se aplicarán anualmente y se espera que sus resultados sean elementos de reflexión para los docentes y directivos que faciliten la identificación de acciones para apoyar a los alumnos, tanto individualmente como en grupo. Se administrarán en cuarto grado pues éste representa un momento de inflexión en el que los alumnos inician la segunda mitad de su educación primaria —en muchos casos en la misma escuela—, lo cual ofrece una oportunidad para redefinir, en caso de ser necesario, el curso del trabajo escolar.

Modalidades de evaluación y esquema de aplicación

En síntesis, los momentos y ciclos de aplicación mencionados en los párrafos anteriores son los siguientes: 1) las ELSN se realizarán cada cuatro años al final del ciclo escolar, empezando en 2015 con sexto de primaria y tercero de secundaria; en 2017 se hará lo mismo con tercero de preescolar y el último grado de educación media superior; 2) las ELCE, se realizarán cada año, al final del ciclo escolar, a partir de 2015 en sexto de primaria, tercero de secundaria y último grado de educación media superior; 3) las EDC se harán cada año a partir de 2015 al inicio de cuarto grado de primaria. El anexo 1 describe brevemente el procedimiento previsto de aplicación para cada una de las tres modalidades.

El siguiente esquema resume el plan de aplicaciones:

Esquema de aplicación

Grado	Año					Revisión del esquema	2020
	2015	2016	2017	2018	2019		
3° de Preescolar			ELSEN				
4° de Primaria	EDC	EDC	EDC	EDC	EDC		EDC
6° de Primaria	ELSEN	ELCE	ELCE	ELCE	ELSEN		ELCE
	ELCE				ELCE		
3° de Secundaria	ELSEN	ELCE	ELCE	ELCE	ELSEN		ELCE
	ELCE				ELCE		
Último grado de media superior	ELCE	ELCE	ELSEN ELCE	ELCE	ELCE		ELCE

ELSEN: Evaluación del logro referida al Sistema Educativo Nacional.
ELCE: Evaluación del logro referida a los centros escolares.
EDC: Evaluación diagnóstica censal.

Independientemente de las revisiones y ajustes que se harán a PLANEA en caso de que se requiera, un momento específico de revisión a detalle de lo alcanzado, para mejorar su diseño e implementación será el año 2019, momento en que se habrá completado un ciclo de aplicaciones. En el apartado 5 de este documento se ofrecen detalles al respecto.

Población objetivo

La población objetivo de las evaluaciones con instrumentos de aplicación externa (del logro referidas al SEN y de logro referidas a los centros escolares) está constituida por los alumnos que terminan el tercer grado de preescolar, el sexto de primaria, el tercero de secundaria y el último grado de educación media superior. Evaluar a los estudiantes al finalizar un nivel ofrece un buen indicador de la eficacia del proceso educativo en su conjunto, reconociendo los logros de los alumnos a lo largo de varios años de trabajo en los que se van integrando diferentes aprendizajes que conforman una red compleja de conocimientos, habilidades y competencias. Además, puede dar respuestas a nivel de la comunidad escolar, a interrogantes tales como: ¿cuáles son los aprendizajes que están alcanzando nuestros alumnos y cuáles no? ¿Qué factores del contexto familiar y económico de los alumnos están relacionados con los resultados?

Para la EDC, a cargo de las escuelas, la población objetivo se constituye por todos los estudiantes que inician el cuarto grado de primaria. La evaluación de los estudiantes al inicio de un ciclo escolar ofrece a los docentes un panorama tanto de los aprendizajes que sus alumnos han adquirido como los que no alcanzaron en cursos anteriores, para buscar estrategias remediales y planear los aprendizajes subsecuentes.

Evaluación de aprendizajes clave

En México existe un plan de estudios para la educación básica de observancia nacional, en el que la autoridad educativa indica lo que deben aprender los alumnos de todo el país.⁹ En educación media superior se cuenta con un Marco Curricular Común (MCC) que define competencias básicas que deben desarrollar todos los egresados de ese nivel,¹⁰ con independencia del plan de estudios que cursen.

Los instrumentos para la evaluación de aprendizajes de PLANEA se diseñan a partir de las habilidades y los conocimientos que el plan de estudios nacional, en el caso de educación básica, y el MMC en el caso de educación media superior, buscan desarrollar, identificando en ambos casos los aprendizajes clave para su evaluación. Dichos aprendizajes clave son relativamente estables en el tiempo, relevantes para el dominio de los conocimientos y habilidades del campo formativo correspondiente, y facilitadores en la adquisición de nuevos aprendizajes. PLANEA evaluará los aprendizajes clave de los campos de formación relacionados con Lenguaje y Comunicación y Matemáticas, que son herramientas esenciales para

⁹ Acuerdo 592 de la SEP (2011).

¹⁰ Acuerdo 444 de la SEP (2008).

el desarrollo del aprendizaje de otras áreas del conocimiento, y buenos indicadores de los resultados educativos en general. En 2017 se incorporará la evaluación de aprendizajes de Ciencias Naturales y Formación Cívica y Ética, en la modalidad de ELSÉN.

Evaluación de habilidades socioafectivas

Se denominan aprendizajes socioafectivos aquellos orientados a la formación social, valoral y emocional de los educandos.¹¹ El currículo nacional de educación básica vigente y el MCC para la educación media superior incluyen este tipo de elementos de manera transversal al abordar temas de relevancia social como la atención a la diversidad, la equidad de género, la educación para la salud, la prevención de la violencia escolar, la educación para la paz y los derechos humanos,¹² entre otros.

Las habilidades socioafectivas son determinantes para tener éxito en la vida escolar y en la adultez, ya que interactúan con los aprendizajes curriculares para ayudar a planear, tomar decisiones y resolver problemas, además resultan fundamentales para el desarrollo y bienestar de las personas. PLANEA incursionará de manera gradual en distintos aspectos de este ámbito de la formación de los educandos.

En la primera aplicación en educación básica, en junio de 2015, se incluirá un acercamiento a la evaluación de aprendizajes relacionados con la convivencia escolar a partir de dos elementos centrales: su anclaje en lo cotidiano, es decir, en el reconocimiento del enorme potencial formativo que tiene el acontecer cotidiano en las escuelas, y la perspectiva de lo público, esto es, la consideración sobre la importancia que tienen estas interacciones diarias en la escuela como formación para desempeños futuros en la vida ciudadana y social (Fierro, Carbajal y Martínez-Parente, 2010).

Una buena convivencia escolar “supone la construcción de una experiencia escolar formativa para desarrollar valores, actitudes y habilidades socioemocionales y éticas que sustentan una convivencia social donde todos participan, comparten y se desarrollan plenamente” (UNESCO, 2008).

El Plan de estudios 2011 de la SEP reconoce la importancia de desarrollar en los tres niveles de educación básica y a lo largo de la vida, una serie de competencias entre las que se encuentran las relacionadas con la convivencia. “Su desarrollo requiere: empatía, relacionarse armónicamente con otros y la naturaleza; ser asertivo; trabajar de manera colaborativa; tomar acuerdos y negociar con otros; crecer con los demás; reconocer y valorar la diversidad social, cultural y lingüística”.¹³ Por ello, tanto en sexto de primaria como en tercero de secundaria se aplicará a los estudiantes un cuestionario que explora seis aspectos de las habilidades para la convivencia: competencia social escolar, conducta prosocial, relación

¹¹ Rosen, J. A., Glennie, E. J., Dalton B. W., Lennon, J. M., y Bozick, R. N. (2010). *Noncognitive skills in the classroom: New perspectives on educational research*. Research Triangle Park, NC: RTI International. Recuperado de <http://www.rti.org/pubs/bk-0004-1009-rosen.pdf>

¹² Cfc. SEP (2011). Acuerdo 592, p. 28.

¹³ SEP (2011). Plan de estudios 2011. Educación básica, p.38.

positiva con pares, prácticas que obstaculizan la convivencia, acoso escolar y manejo pacífico de conflictos. En el Anexo 2 de este documento se definen dichos aspectos.

En la primera aplicación de PLANEA para la educación media superior, en marzo de 2015, se incluyeron escalas de habilidades socio afectivas sobre manejo de estrés, empatía, toma de decisiones y acoso escolar, relacionadas con algunas de las competencias genéricas enunciadas en el MMC para la educación media superior, particularmente con las competencias “Se auto determina y cuida de sí”; “Trabaja en forma colaborativa” y “Participa con responsabilidad en la sociedad”¹⁴

En ambos casos, es la primera vez que se aplican escalas de esta naturaleza a los estudiantes de todas las escuelas del país. La información sobre habilidades para la convivencia tiene sentido por sí misma, independientemente de la vinculación que tenga con otros aprendizajes. Por ello, los análisis que se realicen buscarán dar cuenta tanto de los aprendizajes para la convivencia en diferentes subpoblaciones, como de su relación con los resultados de las pruebas y con los diversos contextos.

La contextualización de las evaluaciones

Muchos factores y circunstancias influyen en el logro educativo de los estudiantes. Por una parte, se encuentra la historia personal y familiar que incluye las condiciones socioeconómicas, acceso a bienes culturales, condiciones para el estudio, situación laboral de la familia, pertenencia étnica, etcétera. Por otra parte, están las características de la escuela en términos de infraestructura, organización, trayectorias y habilidades desarrolladas por sus profesores, entre otras. También influyen las circunstancias que rodean a la escuela, tanto ambientales (como niveles de ruido, elementos climáticos, situaciones de emergencia), y sociales (organización con la comunidad, niveles de bienestar, redes con otras escuelas, violencia, delincuencia, drogadicción), como políticas (acceso a programas de apoyo educativo o de bienestar social, condiciones de conflicto y polarización).

Para contextualizar las diferencias en los resultados de aprendizaje, tanto a nivel de sistema educativo como de plantel escolar, es necesario considerar los elementos del entorno personal, familiar y escolar, incorporando una perspectiva que dé cuenta de los apoyos y los obstáculos existentes en la comunidad educativa para lograr los aprendizajes de los estudiantes. Estos elementos no deberán utilizarse como justificación de resultados bajos, sino como componentes de una reflexión que valore el esfuerzo de estudiantes, docentes y directivos, tomando en cuenta los factores internos y externos que hayan dificultado o facilitado la adquisición de aprendizajes, y que permitan identificar lo que puede mejorarse.

Por esta razón, junto con los instrumentos de las modalidades de evaluación del logro referidas al SEN y a los centros escolares, se aplicarán cuestionarios de contexto que permitirán analizar algunos de los elementos mencionados. En el anexo 3 se presenta la estructura de los cuestionarios que serán aplicados en 2015. Estos instrumentos podrán

¹⁴ SEP (2008). Acuerdo número 444 por el que se establecen las competencias que constituyen el marco curricular común del Sistema Nacional de Bachillerato. Artículos 3 y 4.

sufrir modificaciones a lo largo del tiempo, a partir de temas emergentes que se considere necesario tomar en cuenta. Sin embargo, existe la intención de conservar una base uniforme de información de contexto a lo largo del tiempo, que ayude a comprender los resultados comparativos.

Por sus características, los cuestionarios de contexto no pueden dar cuenta con precisión de todos los elementos recién enumerados por lo que se buscará incorporar fuentes secundarias de información (por ejemplo, bases de datos nacionales como la que se genera a partir del Formato 911, o del proyecto Evaluación de las Condiciones Básicas para la Enseñanza y el Aprendizaje [ECEA] del INEE, que contienen información relevante sobre las escuelas y sus alrededores) que permitan completar el panorama que se ofrecerá tanto en los informes de resultados nacionales como en los escolares.

Participación de diferentes grupos de expertos en el diseño y elaboración de los instrumentos de evaluación

Un elemento metodológico fundamental al diseñar y desarrollar los instrumentos de evaluación que se utilizan para PLANEA es la toma de decisiones colegiada, a partir de la reflexión de grupos de especialistas coordinados por el INEE.

Los estándares internacionales sobre prácticas justas en evaluación señalan la importancia de asegurar que las pruebas eviten obstáculos para que algunos grupos de alumnos demuestren apropiadamente sus niveles de logro.¹⁵ Este tema es muy importante en el contexto de las evaluaciones nacionales de un país tan diverso como el nuestro, y por ello los instrumentos de PLANEA se desarrollan a lo largo de todas sus fases por grupos de trabajo que incorporan tanto a personas con experiencia en el ámbito educativo como a expertos en diversidad cultural y lingüística; ambos grupos participan en las discusiones y en la toma de decisiones, a fin de lograr que las pruebas den a todos los evaluados la misma oportunidad de demostrar sus aprendizajes sin obstáculos.

La preparación y aplicación de las nuevas pruebas se realiza en las siguientes etapas:

- **Selección de contenidos a evaluar:** Se reúnen Comités Académicos por asignatura y nivel para determinar los aprendizajes clave que deberán evaluarse. Estos comités están conformados por autoridades educativas relacionadas con el currículo y con materiales educativos, autores de libros de texto y especialistas en didáctica de cada campo formativo.
- **Determinación de la manera de evaluar los contenidos:** Se reúnen Comités de Especificaciones por asignatura y nivel para precisar y operacionalizar los contenidos a evaluar a partir de las definiciones construidas por los Comités Académicos. Estos comités están conformados por expertos en didáctica de los campos formativos evaluados, y por especialistas en contenidos escolares. En su capacitación

¹⁵ Cfc. AERA, APA, NCME (2014). *Standards for Educational and Psychological Testing*. Washington, DC: American Educational Research Association, pp. 49-50.

se incluyen elementos para considerar la diversidad cultural de nuestro país en el desarrollo de especificaciones; entre los revisores de las especificaciones se encuentran especialistas en educación multicultural o en diversidad.

- **Elaboración de los reactivos:** A partir de la formulación de especificaciones realizada por los Comités descritos en el párrafo anterior, se reúnen especialistas en los contenidos escolares para desarrollar reactivos técnicamente correctos. Cada reactivo es revisado además por otros dos especialistas uno de los cuales, al menos, es especialista en diversidad cultural en educación, y por el coordinador de la prueba, que es experto en la didáctica del campo formativo a evaluar.
- **Validación de los reactivos:** Como se señala en los *Criterios técnicos para el desarrollo y uso de instrumentos* del INEE, resulta fundamental que los reactivos sean validados por un grupo independiente de especialistas que determinen en qué medida los reactivos atienden al contenido que deben medir, y hagan observaciones sobre diferentes tipos de sesgo que las versiones iniciales pudieran tener. Para llevar a cabo estas tareas, se convoca a profesoras y profesores en servicio de los grados escolares correspondientes a las pruebas, de las 32 entidades de la república, de distintas modalidades educativas, distintos grupos étnicos, entidades y condiciones socioeconómicas para que analicen de manera crítica y sistemática todos los reactivos formulados por el comité anterior y verifiquen, entre otros aspectos, que los reactivos sean adecuados para estudiantes que viven en distintos ambientes y circunstancias.
- **Piloteo de los reactivos y desarrollo de versiones finales de las pruebas:** Todos los reactivos redactados, revisados y validados se someten a pruebas piloto antes de incorporarse en aplicaciones definitivas; se desarrollan análisis psicométricos y se determina cuáles reactivos son adecuados para las aplicaciones finales. Las pruebas piloto siempre incluyen escuelas rurales e indígenas como parte de sus muestras, y la información psicométrica sobre la calidad de los reactivos que se obtienen de estos grupos vulnerables se contrasta con el resto de la muestra, prefiriéndose siempre los reactivos que tienen indicadores adecuados tanto para el total de alumnos participantes en el piloteo como para los alumnos de escuelas en contextos desfavorecidos. Como resultado de este proceso, se desarrollan las versiones finales de las pruebas que son revisadas editorialmente para asegurar la comprensión de las preguntas.

IV. Precisiones sobre las modalidades de PLANEA

Como ya se mencionó, PLANEA presenta tres modalidades de evaluación. Esto obedece a que cada una de ellas resulta más adecuada para obtener distintos tipos de información, y por lo tanto, para alcanzar algunos de los propósitos enunciados: la modalidad ELSEN incluye una amplia cantidad de contenidos a evaluar, pero cada alumno sólo responderá una parte de las pruebas, lo cual es adecuado cuando se quiere obtener información de grandes grupos de individuos. En la modalidad ELCE, los alumnos de todas las escuelas responden un mismo grupo de preguntas, considerando una cantidad moderada de contenidos, lo cual es conveniente cuando se quiere obtener información a nivel de pequeños grupos o de

individuos. Finalmente, los instrumentos de la EDC son administrados y analizados para su uso por los propios docentes al inicio del ciclo escolar, porque en este caso se trata de información que sirve para conocer los aprendizajes del grado escolar anterior que los alumnos han alcanzado y los que aún no dominan, tanto grupal como individualmente.

Evaluación de logro referida al Sistema Educativo Nacional (ELSEN)

Se aplicarán a muestras representativas de estudiantes del país de todos los grados terminales de la educación básica, en ciclos de cuatro años. Esta frecuencia es suficiente para monitorear cambios a nivel del sistema educativo que no suelen darse en plazos cortos. Tienen el propósito específico de ofrecer al SEN en su conjunto elementos de retroalimentación y aportar información a los tomadores de decisiones de política educativa, y para la rendición de cuentas. Las pruebas son diseñadas, desarrolladas, aplicadas y analizadas por el INEE. Esta modalidad se incorporará en educación media superior a partir de 2017.

Diseño matricial

Con la intención de evaluar un conjunto amplio de contenidos del currículo, se implementa un diseño denominado matricial en el que cada alumno evaluado contesta solamente un subconjunto del total de reactivos que integran la evaluación. A partir del universo de reactivos se construyen formas de examen relacionadas entre sí que al ser aplicadas a los alumnos, permiten conocer lo que saben de ese conjunto amplio de contenidos.

Los reactivos de cada uno de estos exámenes, se agrupan en seis conjuntos llamados bloques. Los bloques de reactivos tienen las siguientes características:

- Tienen una dificultad promedio similar.
- La distribución de dificultades también es similar.
- Incluyen reactivos de cada uno de los niveles de desempeño de cada categoría de la evaluación (dimensiones o ejes temáticos).
- Tienen la misma cantidad de reactivos (de 25 a 30).

Para su administración a los alumnos, los seis bloques de reactivos de una evaluación se agrupan en seis formas de examen, como se muestran a continuación:

Forma	Bloques	
1	A	B
2	B	C
3	C	D
4	D	E
5	E	F
6	F	A

Cada uno de los alumnos evaluados resuelve una forma de examen. La forma 1 está conformada por los bloques A y B, la forma 2 por los bloques B y C, y así sucesivamente. Todas las formas de examen tienen entre 50 y 60 reactivos.

Se dice que el diseño es balanceado porque cada uno de los bloques se administra al inicio y al final de alguna de las formas de examen. Asimismo, el diseño es incompleto porque cada alumno solamente resuelve una parte de la evaluación.

El diseño matricial permite obtener las puntuaciones en una escala en común por la manera en que se eslabonan los bloques de reactivos. Así por ejemplo el bloque A se incluye en la forma 1 y en la forma 6, estableciendo un vínculo con el bloque B y el bloque F.

Por el modo en que se construyen, cada forma de examen es una muestra representativa de la evaluación completa por lo que es factible hacer inferencias del desempeño de los evaluados respecto a toda la evaluación; es decir, un alumno no necesita responder a todos los reactivos de la evaluación para valorar su nivel de desempeño en la escala completa, la cual se construye con base en todos los reactivos que la conforman.

En el anexo 4 se presentan elementos centrales del diseño de la muestra, incluyendo los dominios de estudio que son considerados para esta modalidad. Es de destacarse que se incluyen los cursos comunitarios administrados por el Consejo Nacional de Fomento Educativo (CONAFE), lo cual requiere un esfuerzo especial debido a las difíciles condiciones de acceso por razones tanto geográficas como sociales.¹⁶ Conviene adelantar que para las aplicaciones de la segunda modalidad que se refieren a continuación, los análisis realizados junto con la autoridad educativa mostraron barreras importantes para su aplicación en los cursos comunitarios a través de evaluadores externos, por lo que su administración se depositará en las mismas autoridades de CONAFE, dando además tiempos extendidos para la aplicación, posteriores a la del resto de los centros escolares.

Evaluación del logro referida a los centros escolares (ELCE)

Son versiones derivadas de los instrumentos matriciales de la primera modalidad. La SEP, en coordinación con las autoridades educativas estatales, las aplicará anualmente a estudiantes de todas las escuelas del país en los grados terminales de primaria, secundaria y bachillerato. Tienen el propósito específico de ofrecer a cada escuela elementos de retroalimentación para su mejora. Los resultados de estas pruebas dan información de lo que logró el centro escolar con respecto al aprendizaje de sus estudiantes en seis años (para el caso de primaria); indican lo que los alumnos logran y no logran aprender; ayudan a identificar las líneas curriculares que la comunidad escolar debe fortalecer desde el primer grado de primaria, y facilitan elementos para identificar acciones que pueden implementarse a fin de lograr los aprendizajes esperados; propician que el Consejo Técnico y los supervisores centren su atención en el aprendizaje de los estudiantes; facilitan el monitoreo de la

¹⁶ El documento PLANEA 2015. Sexto Grado de Primaria y Tercer Grado de Secundaria. Diseño Muestral, elaborado por la Dirección General de Medición y Tratamiento de Datos del Instituto detalla las características técnicas del diseño matricial, de la muestra, y su justificación.

eficacia de cada escuela a través del tiempo; permiten que un centro escolar se compare con los resultados agregados de otros similares; y promueve el interés y colaboración del colegiado para lograr mejores resultados en los siguientes años.

Los instrumentos de estas evaluaciones y los de las referidas al SEN estarán alineadas entre sí, partirán de un mismo análisis curricular y compartirán contenidos, de manera que ambas reportarán información sobre los mismos constructos o dimensiones educativas, utilizando unidades de medida congruentes que permitirán examinar tanto los resultados agregados a nivel nacional como los de una escuela en particular. Sin embargo, por exigencia de los métodos de aplicación, los instrumentos de las ELCE serán menos extensos que los de las ELSN. En los planteles donde el INEE aplique éstas últimas no se aplicarán las primeras; sin embargo, los resultados por escuela se entregarán exactamente de la misma manera que en el resto de los planteles escolares.

Las pruebas de las ELCE serán aplicadas por las autoridades educativas y sus resultados serán analizados en conjunto por el INEE y las autoridades educativas. La SEP pondrá a disposición de cada escuela sus resultados, acompañados de elementos que permitan contextualizarlos adecuadamente. Los alumnos que no queden seleccionados en la muestra tendrán la opción de presentar el examen en línea.

Esta modalidad no se utilizará en preescolar pues las pruebas que se administran en este nivel, al requerir un trabajo del aplicador con cada niño, no permiten un diseño que implique su aplicación en todas las escuelas del país.

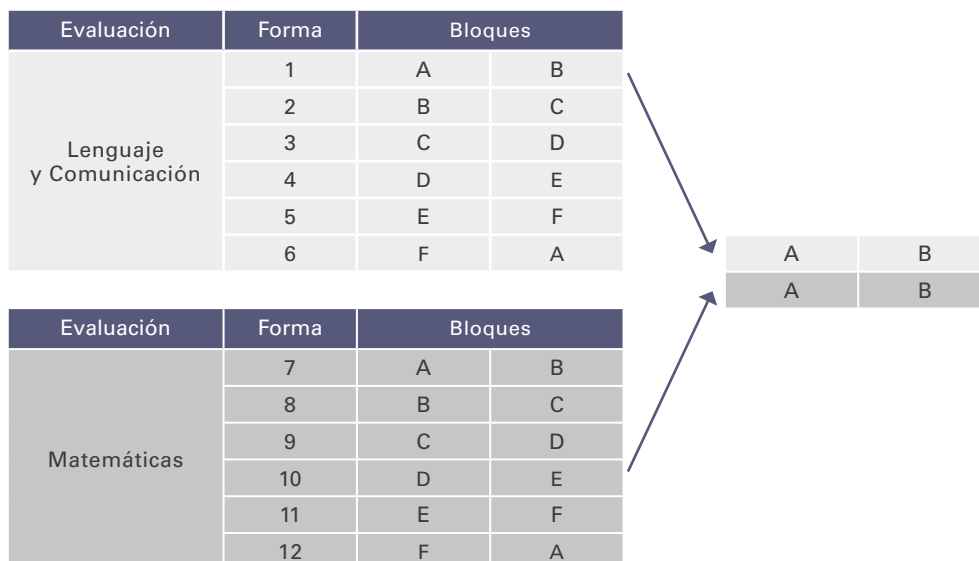
Instrumentos para la evaluación del logro referida a los centros escolares

Para llevar a cabo los operativos de esta modalidad se utiliza una de las formas de examen del diseño matricial presentado anteriormente. En principio, es indistinta la forma de examen que se seleccione debido a la manera en que fueron construidas.

A continuación se presenta el esquema que ejemplifica la conformación de los instrumentos de Lenguaje y Comunicación y de Matemáticas para la ELCE, a partir de la selección de la forma 1 y la forma 7 de los instrumentos de la ELSN.

Evaluación del logro referida al Sistema Educativo Nacional.
Arreglo de formas matriciales

Evaluación del logro referida a los centros escolares
Forma única a aplicar



Nota: Se han utilizado las mismas letras para los bloques de la evaluación de Lenguaje y Comunicación y la de Matemáticas solamente para efectos del ejemplo y no significa que sean iguales para ambas evaluaciones.

Renovación y mantenimiento de las evaluaciones

Una vez que los instrumentos de la ELCE son divulgados, las formas de examen que se utilizaron no volverán a ocuparse, evitando así sesgos en los resultados; por ello se implementará una estrategia de renovación y mantenimiento de las evaluaciones, que permita hacer inferencias válidas y confiables a partir de los resultados tanto de las evaluaciones de logro referidas al SEN como de las referidas a los centros escolares.

A partir de 2016, cada año se incluirán en el diseño matricial de los instrumentos para la evaluación del sistema bloques de reactivos que reemplazarán a los utilizados en las ELCE.

Los nuevos bloques que se incorporen, deberán calibrarse y ubicarse en la misma métrica definida en 2015. Para ello, las formas de examen del diseño matricial se aplicarán a una muestra de alumnos de 150 escuelas aproximadamente. Este procedimiento se llevará a cabo en los años en que no se apliquen las evaluaciones de logro referidas al SEN: 2016, 2017 y 2018. Así, en 2016 se deberán calibrar los bloques A' y B' que sustituyan a los bloques A y B utilizados para la ELCE de 2015.

A continuación se muestra el esquema para la renovación y el mantenimiento de las evaluaciones. Los bloques de reactivos a utilizar en las Evaluaciones del logro referidas a los centros escolares (ELCE) están sombreados. Las evaluaciones del logro referidas al SEN se señalan como ELSN.

ELSEN								ELSEN	
ELCE		ELCE		ELCE		ELCE		ELCE	
2015		2016		2017		2018		2019	
A	B	A'	B'	A'	B'	A'	B'	A''	B''
B	C	B'	C	B'	C'	B'	C'	B''	C'
C	D	C	D	C'	D'	C'	D'	C'	D'
D	E	D	E	D'	E	D'	E'	D'	E'
E	F	E	F	E	F	E'	F'	E'	F'
F	A	F	A'	F	A'	F'	A'	F'	A''

Bloques nuevos que se incorporarán

A'	C'	E'	A''
B'	D'	F'	B''

Para la aplicación 2016 por ejemplo, se deberán incorporar al diseño matricial los bloques A' y B'; y se utilizarán para la evaluación referida a los centros escolares los bloques C y D. Puede observarse que para el año 2018, todos los bloques del diseño matricial habrán sido renovados respecto a los utilizados en 2015; este esquema de renovación y mantenimiento permitirá incorporar paulatinamente mejoras en los instrumentos.

Mecanismos de control de las aplicaciones relacionadas con la Evaluación del logro referida a centros escolares

Las pruebas de aprendizaje en las que puede identificarse a los individuos o a las escuelas presentan a menudo un aumento en las puntuaciones que no corresponde con los niveles reales de los alumnos, exagerándolos. "Cuando se inflan los resultados, muchas de las conclusiones más importantes que la gente desprende de ellos resultan equivocadas. Esto afecta a los estudiantes y, en ocasiones, a los maestros".¹⁷ Aunque este fenómeno es más frecuente en evaluaciones de alto impacto (aquellas cuyos resultados afectan directamente a los individuos, por ejemplo, las que sirven a concursos de ingreso o para dar incentivos), el simple deseo de los miembros de una comunidad escolar por obtener buenos resultados por razones de prestigio puede causar su inflación. Por esta razón se han considerado tres mecanismos de control para la aplicación de los instrumentos de la evaluación referida a los centros escolares:

1. El INEE desarrolla lineamientos específicos para la aplicación de las pruebas, entre los que se indica que los aplicadores de las pruebas en un plantel no deben ser parte de esa comunidad escolar. También se dan indicaciones relativas al aseguramiento de condiciones para que los alumnos puedan mostrar lo que saben sin problemas que distorsionen los resultados.

¹⁷ Koretz, D. (2010). *El ABC de la evaluación educativa* (Measuring Up). México: CENEVAL, p. 275.

2. El INEE supervisa la aplicación y los procesos asociados a la misma. Con diferente intensidad y en diversos momentos, el INEE acompaña los procesos de preparación, aplicación y procesamiento de la información de las pruebas de la evaluación referida a los centros escolares, con la intención de verificar que el trabajo de campo tenga las condiciones que permitan asegurar una buena confiabilidad en los resultados.
3. El INEE realiza una verificación estadística de la consistencia de los resultados encontrados entre las aplicaciones de las ELCE y las de las ELSN. La intención de esta verificación es asegurar la congruencia entre los resultados de ambas modalidades, e identificar los casos en que no la haya. A partir de la suposición razonable de que una muestra más pequeña probablemente será mejor controlada que una muestra mayor, se contrasta la distribución de resultados de la aplicación de la muestra controlada por el INEE con la de la aplicación al resto de las escuelas controlada por la SEP. Se determina si las diferencias encontradas son significativas, en cuyo caso se hacen otros análisis para establecer en dónde se encuentran las discrepancias, y si éstas se deben a malas prácticas tales como patrones de copia —que pueden referirse a diferentes formas de hacer llegar las respuestas correctas a los estudiantes, y no solamente a copia entre dos o más alumnos—; y se establecen mecanismos para detectar estos patrones a diferentes niveles de desagregación. En el anexo 5 se presenta una síntesis del esquema de verificación estadística de los resultados.¹⁸

La aplicación de estas medidas podrán llevar a que el INEE señale que los resultados de determinada entidad, modalidad educativa o escuela no son confiables, y por lo tanto, a recomendar su uso limitado.

Evaluación diagnóstica censal (EDC)

La tercera modalidad de PLANEA se desarrolla a partir del propósito central de ofrecer información para la mejora de los procesos de enseñanza en los centros escolares. En particular, el diseño de esta modalidad permite apoyar a los docentes frente a grupo con un instrumento que facilite la realización de un diagnóstico de sus alumnos al inicio del ciclo escolar, sobre los conocimientos y habilidades que debieron haber logrado durante los cursos anteriores. Esto favorecerá que los docentes puedan:

1. Ajustar la planeación de la intervención pedagógica a las características y necesidades de sus alumnos.
2. Detectar de manera oportuna a los alumnos que requerirán mayor atención para concluir con éxito el ciclo escolar.
3. Proveer retroalimentación personalizada a los estudiantes.
4. Establecer una línea base con la cual la comunidad escolar podrá identificar la eficacia de la intervención pedagógica.

¹⁸ La Dirección General para la Integración y Análisis de la Información desarrolla un documento extenso que explica con detalle la estrategia de control estadístico.

Muchos docentes frente a grupo tienen herramientas conceptuales y metodológicas adecuadas para realizar un diagnóstico que les permita llevar a cabo las tareas mencionadas; sin embargo, el instrumento estandarizado que se ofrece a todos los docentes de cuarto grado del país presenta algunas características adicionales que se espera sean de utilidad:

- La selección de contenidos para las pruebas es llevada a cabo colegiadamente por especialistas en el currículum de educación básica, por lo que los instrumentos presentan equilibrio en cuanto a la diversidad de temáticas consideradas.
- Profesores de diferentes aulas, e incluso de diferentes escuelas, tendrán elementos en común para la reflexión y para el diálogo entre pares.
- Junto con los instrumentos se proporciona a los docentes una guía que incluye los propósitos de esta evaluación, la manera de aplicarla y de analizar los resultados; con ello se propicia una reflexión amplia sobre lo que el docente puede hacer considerando los aprendizajes con los que llegan los estudiantes a su curso.

Esta modalidad de PLANEA será operada por la SEP. El INEE proporcionará tablas de contenidos, análisis curriculares, el banco de reactivos inicial y su valoración psicométrica para el inicio de los trabajos, acompañamiento técnico a lo largo de su desarrollo, y en su caso, la aprobación de las pruebas resultantes. Ambas instancias harán un análisis y selección de los reactivos más adecuados para los propósitos de esta modalidad y la SEP conformará y distribuirá los instrumentos a los docentes de cuarto grado de primaria a inicios del ciclo escolar. Éstos se presentarán como parte de la actividad pedagógica que realiza el docente, con la intención de que los resultados que obtenga se utilicen sólo al interior de la escuela; no se divulgarán al exterior y no servirán para la rendición de cuentas. El proyecto iniciará en el ciclo 2015-2016 con pruebas de Lenguaje y comunicación y Matemáticas y después abordará otros campos formativos. Además la SEP ofrecerá instructivos para calificar e interpretar los resultados, y un *software* de apoyo opcional para la generación de reportes para el docente y la comunidad escolar.

Participación de diferentes actores en las modalidades de PLANEA

El siguiente cuadro resume la participación de diferentes instancias en cada modalidad de PLANEA:

	Evaluación referida al SEN (ELSEN)	Evaluación referida a los centros escolares (ELCE)	Evaluación diagnóstica censal (EDC)
¿Quién desarrolla los instrumentos?	El INEE	El INEE	La SEP y el INEE
¿Quién los aplica?	El INEE	La SEP	El docente
¿Quién los califica y analiza?	El INEE	La SEP y el INEE	El docente
¿Quién utiliza los resultados en primera instancia?	Las autoridades educativas nacionales y estatales. La sociedad. El INEE	La comunidad escolar	El docente. La comunidad escolar

V. Difusión y uso de resultados

Una de las preocupaciones fundamentales de PLANEA son las necesidades de los distintos tipos de usuarios de todos los niveles del sistema educativo que se espera aprovechen los resultados de la evaluación para enriquecer sus decisiones. En tal sentido, se acompañará el proceso de diseño y aplicación de los instrumentos con un trabajo sistemático de identificación de expectativas, necesidades y usos potenciales de la información, atendiendo a los retos que suponen los distintos procesos que enfrenta cada tipo de actor al realizar su quehacer educativo —desde el diseño de políticas, planeación, gestión, monitoreo y rendición de cuentas, hasta la mejora de la enseñanza en los centros escolares.

PLANEA se ha diseñado en el contexto de un nuevo marco legal para el INEE y otros actores educativos, que institucionaliza canales de comunicación directa y constante con las autoridades educativas federales, estatales y otros participantes del proceso educativo, de los sectores social, público y privado; ello ofrece oportunidades constantes para recuperar una amplia variedad de voces respecto a las necesidades de información y para poner a consideración de diversos grupos los planes de difusión y uso de los resultados.

Esta interacción facilita el desarrollo de un sistema para diseminar los resultados y alentar el uso de la información, teniendo en cuenta la pertinencia de los contenidos, la utilidad de los resultados —tanto en lo que se refiere a formatos y niveles de desagregación como a la oportunidad en la difusión— y la adaptación de medios y lenguajes para hacerla comprensible y utilizable para las diversas audiencias en cuestión.

A partir del diálogo con varios de los órganos colegiados del Instituto, así como algunos de sus Consejos Técnicos Especializados, se han alcanzado las siguientes definiciones respecto a la difusión y uso de los resultados de PLANEA:

- Los resultados de las evaluaciones se presentarán siempre de manera contextualizada, considerando tanto los factores de orden socioeconómico y cultural, como las condiciones del centro de trabajo.
- Se señalarán los casos en que los resultados de las evaluaciones no tengan garantía de ser confiables, recomendando su uso limitado. Esta situación puede darse, por ejemplo, si el número de alumnos participantes de un centro escolar no es suficiente, si se encuentran patrones de copia en una cantidad relevante de pruebas, o si las condiciones de aplicación no son adecuadas.
- Las unidades de análisis para las ELSN son las entidades y los tipos de servicio, mientras que para las ELCE, éstos serán la unidad de análisis fundamental. Para las evaluaciones diagnósticas censales la unidad de análisis es el aula.
- Como se ha mencionado reiteradamente, los resultados de estas evaluaciones no deberían utilizarse para juzgar el desempeño de los docentes, realizar *rankings* de escuelas, justificar procesos punitivos u otros de control administrativo sobre estudiantes, docentes o escuelas, puesto que no están diseñados para ello ni aportan información suficiente para realizar estas tareas con justicia.

A partir de los resultados de las evaluaciones con aplicadores externos (ELSEN y ELCE) se consideran los siguientes documentos informativos para distintas audiencias:

- 1. Reporte de Evaluación del logro referida al SEN.** Informe breve a nivel de grandes estratos, con apoyos gráficos y textos breves. Lo prepara el INEE y se entrega pocos meses después de la aplicación. Además de hacerse público, este reporte se presentará en la Conferencia del Sistema Nacional de Evaluación Educativa, órgano colegiado integrado por la Junta de Gobierno del INEE, por los Subsecretarios y representantes de la SEP y por los Secretarios de Educación u homólogos de las entidades federativas. El reporte debe incluir un apartado específico que muestre las diferencias en los resultados entre los grupos sociales más favorecidos y los más desfavorecidos, de manera que sean muy claras las brechas que hay que eliminar para alcanzar la equidad en los aprendizajes.
- 2. Reportes de la Evaluación del logro referida a centros escolares.** Son preparados por la SEP con acompañamiento del INEE para la definición de la información a incorporar y la manera de presentarla, y tienen diferentes niveles de desagregación:
 - a. Dirigido a las autoridades locales: datos de escuelas agregados, y desagregados por región o municipio.
 - b. Dirigido a supervisores escolares: datos de escuelas agregados, y desagregados por zona escolar.
 - c. Dirigido a la comunidad escolar: datos de su escuela y de escuelas similares en su entidad (por ejemplo, modalidad, tamaño de localidad, tipo de organización, etcétera). El INEE y la SEP alentarán que estos informes se utilicen tanto en los Consejos Técnicos Escolares como en los Consejos Escolares de Participación Social.
- 3. Bases de datos.** Se pondrán a disposición de universidades, asociaciones, y público en general, de manera que sea posible realizar análisis complementarios a los que la SEP y el INEE aporten. Este último documentará y difundirá las bases de datos de las ELSEN y la SEP las de las aquéllas referidas a los centros escolares.
- 4. Informe nacional** sobre el estado que guarda la educación con respecto al logro de aprendizajes.
- 5. Otros informes temáticos.** La SEP, el Instituto y otras instancias podrán desarrollar informes centrados en alguna temática en particular, apoyados en información empírica extraída de los resultados de las evaluaciones, por ejemplo, sobre los resultados de los estudiantes en condiciones de extraedad con respecto a los estudiantes con un trayecto escolar regular, sobre los resultados en las escuelas multigrado, o sobre la identificación de características escolares y familiares de grupos con menores niveles de logro.

Reportes de la evaluación del logro referida a los centros escolares

Los reportes de la ELCE que se entregan a la escuela son un elemento central para la devolución de resultados. Que la comunidad escolar conozca sus resultados, dialogue sobre ellos y extraiga conclusiones que les permita apoyar decisiones y acciones para la mejora es uno de los propósitos centrales de PLANEA. Los reportes incluirán la siguiente información:

- Datos de la escuela y de escuelas similares en su entidad (misma modalidad, elementos sobre el nivel socioeconómico de las familias, si la escuela es de organización completa o no).
- Número de alumnos que aplicaron/número total de alumnos en la escuela.
- Confiabilidad de los resultados (porcentaje de alumnos de la muestra que participaron, detección de copia).
- Tablas o gráficas con porcentaje de alumnos por nivel de logro, por campo formativo (Lenguaje y Comunicación y Matemáticas).
- Condiciones de la oferta educativa en el plantel, variables para la reflexión de la comunidad escolar y elementos sobre el funcionamiento de la escuela.

Junto con estos reportes, se entregará a la comunidad escolar una serie de materiales para facilitar la comprensión e interpretación de los resultados:

- Folleto que explica los propósitos y las características principales de las pruebas PLANEA.
- Descriptores de los niveles de logro, que indican qué es lo que ya dominan los alumnos que se encuentran en cada nivel de logro y qué aprendizajes clave están aún por desarrollar.
- Guía para entender y analizar los reportes para la comunidad escolar (en los Consejos Técnicos Escolares y en los Consejos Escolares de Participación Social), incluyendo preguntas relevantes para reflexionar en las posibilidades de mejora de los aprendizajes desde lo que la escuela puede hacer para todos los alumnos, la atención compensatoria que se puede ofrecer para los que provienen de entornos más vulnerables, y la búsqueda de mejores esquemas de atención a la diversidad. Esta guía puede incluir gráficas o tablas que muestren la relación entre las categorías de contexto y el rendimiento escolar.
- Guía para entender los cambios anuales y los cambios abruptos en los resultados en los resultados (esto es necesario sobre todo en escuelas de menos de 10 alumnos participantes, en que las variaciones de los resultados pueden ser muy pronunciados de un año a otro).
- Pruebas, y revisión de las pruebas, indicando cuáles son las respuestas correctas (y por qué esas son correctas y las demás equivocadas).
- Acceso a página web en donde puedan consultarse los resultados históricos de las escuelas.

La SEP y el INEE, en coordinación con las autoridades educativas estatales desarrollan un esquema de seguimiento que permita conocer los usos que están dando los docentes a los reportes de resultados, y las áreas de oportunidad que se tienen en la difusión de resultados.

VI. Revisión y seguimiento del Plan

Es fundamental contar con estudios que muestren la utilidad de la información obtenida a partir de PLANEA para la mejora educativa. Independientemente de que exista una serie de revisiones parciales que lleven a depurar los métodos y estrategias de trabajo, y de que haya adecuaciones y agregaciones graduales, PLANEA contempla una revisión a profundidad de sus alcances y limitaciones tras el primer ciclo de aplicación completo, durante el año 2019. Esta revisión deberá ser llevada a cabo por especialistas externos al INEE, a quienes se proporcionará toda la información pertinente para que puedan hacer recomendaciones que mejoren las prácticas de desarrollo, aplicación y análisis de las pruebas, y que potencien los usos de sus resultados. Con base en estas recomendaciones el Instituto podría redefinir los propósitos de PLANEA, sus horizontes técnico y logístico y su esquema de aplicaciones. En principio, parece fundamental que en la revisión se respondan preguntas como las siguientes:

- ¿En qué medida los propósitos de PLANEA siguen siendo vigentes?
- ¿Qué tanto las tres modalidades de PLANEA han logrado el cumplimiento de sus propósitos? ¿En qué medida las disposiciones de las tres modalidades son congruentes con los propósitos?
- ¿El esquema de aplicación de PLANEA ha sido parsimonioso o puede conseguirse información suficiente para sus propósitos con menos aplicaciones? O por el contrario ¿las aplicaciones han sido insuficientes para cumplir sus propósitos?
- ¿Deben incorporarse nuevas modalidades a PLANEA? ¿Ha de suprimirse alguna de las existentes? ¿Deben ampliarse o disminuirse las poblaciones atendidas, los grados escolares de aplicación o las áreas formativas evaluadas?
- ¿Se está atendiendo correctamente en el Plan la diversidad cultural del país? ¿Qué modificaciones a las modalidades de evaluación, al esquema de aplicaciones o a los instrumentos deberían hacerse para atenderla de mejor manera?
- ¿Cómo puede aumentarse la precisión y pertinencia de los resultados?
- ¿Cómo puede mejorarse la comunicación y difusión de resultados?
- Los resultados que se presentan ¿dan cuenta correctamente del estado de inequidad educativa? ¿Son suficientes para caracterizar la situación de los diferentes grupos que requieren más atención?
- ¿Cuáles han sido los usos que diferentes tipos de actores educativos han hecho de los resultados de PLANEA? ¿Son adecuados los métodos y formatos de devolución o se requiere cambiarlos o diversificarlos? ¿Cómo puede alentarse un mejor uso de los resultados?
- ¿Cómo puede alentarse una mayor reflexión de las comunidades escolares para el uso de los resultados y los elementos contextuales en pro de identificar y apoyar más a los estudiantes que más lo necesitan?
- ¿Cómo funcionan los mecanismos incorporados en PLANEA para evitar la inflación de resultados? ¿Deben agregarse o suprimirse mecanismos que tengan este fin?
- ¿Existen usos emergentes de los resultados de PLANEA que sean deseables y que por lo tanto puedan incorporarse al esquema?
- ¿Existen usos perniciosos de los resultados de PLANEA que no fueron contemplados ni controlados con el esquema original? ¿Qué elementos deben incorporarse a PLANEA para que estos usos sean evitados?

Anexo 1.

Procedimiento previsto de aplicación

A. Aplicación conjunta de los instrumentos para la Evaluación del logro referida al Sistema Educativo Nacional (ELSEN) y para la Evaluación del logro referida a los centros escolares (ELCE)

Los operativos de las evaluaciones de PLANEA con aplicadores externos (las referidas al SEN y las referidas a los centros escolares) se aplican en dos días:

- Primer día, Lenguaje y Comunicación
- Segundo día, Matemáticas

Evaluación del logro referida al Sistema Educativo Nacional

En los operativos de esta modalidad se selecciona una muestra de alumnos de la escuela, y se forman entre uno y tres grupos de aplicación dependiendo del total de estudiantes; todos los seleccionados resuelven evaluaciones de Lenguaje y Comunicación y de Matemáticas.

En las aulas de aplicación, las personas encargadas de la administración de los instrumentos, reparten las formas de examen en orden ascendente y de manera cíclica, lo que permite asumir que cada forma de examen es aplicada a una población equivalente de alumnos. Por ejemplo, para la evaluación de Matemáticas (formas 7, 8, 9, 10, 11 y 12), un aplicador reparte las formas de examen en el aula de aplicación en la siguiente secuencia: 9, 10, 11, 12, 7, 8, 9, 10, 11, 12, 7, 8, 9, 10, ... y así sucesivamente.

Sobre el *linking* entre PLANEA y EXCALE

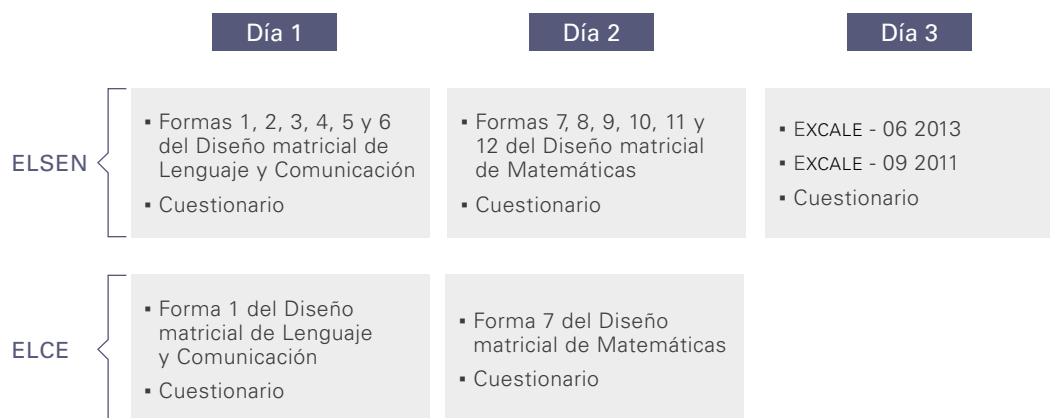
Con el objetivo de definir una línea basal para las evaluaciones de PLANEA y construir la serie histórica de resultados a través del tiempo, se incorporan los resultados obtenidos en las pruebas EXCALE en la aplicación de PLANEA de 2015 y, por única ocasión, se administran las evaluaciones EXCALE —06, 2013 (sexto grado de primaria) y EXCALE— 09, 2012

(tercer grado de secundaria) de Español y Matemáticas a un subconjunto de alumnos de la muestra (aproximadamente 150 escuelas elegidas aleatoriamente) en un tercer día de aplicación. Esto desmotiva a la sociedad a inferir equivocadamente que hay un retroceso en lo que sabemos que aprenden los alumnos; sin embargo, no implica que los niveles de desempeño puedan compararse ya que se miden constructos diferentes. Este procedimiento se conoce como *linking* y es utilizado por instituciones con una larga experiencia en evaluación de logro educativo como NAEP, principalmente para encontrar una regla de correspondencia aproximada entre instrumentos de evaluación que no se basan en las mismas especificaciones técnicas.

Evaluación del logro referida a los centros escolares

Para esta modalidad se selecciona una muestra de alumnos en cada escuela, conformando un máximo de dos grupos de aplicación, y se les administra una de las formas de la ELSN, tanto para Lenguaje y Comunicación como para Matemáticas.

A continuación se presenta un cuadro en donde se resume el procedimiento de aplicación previsto para las dos modalidades. El cuestionario del alumno que se aplica el primer día corresponde a las escalas socioafectivas de habilidades para la convivencia, mientras que el que se aplica el segundo día corresponde a la información de contexto.



Supervisión del INEE en la evaluación del logro referida a los centros escolares

Uno de los mecanismos de control que el INEE implementará para prevenir la inflación de los resultados de manera artificial es la supervisión de la aplicación de las pruebas de la ELCE. El INEE comparte con la SEP sus Criterios y procedimientos para la selección y capacitación de aplicadores que participen en las pruebas del Plan Nacional para las Evaluaciones de los Aprendizajes, en los cuales se detalla el perfil de los aplicadores, los procedimientos que deben seguirse para la selección y capacitación de aplicadores, y en los que se señala que las autoridades educativas y los organismos descentralizados establecerán mecanismos para valorar el trabajo realizado por los aplicadores

a fin de contar con un registro de su desempeño; también se señala que el Instituto podrá supervisar cualquier parte del proceso de selección y capacitación de los aplicadores con la finalidad de garantizar su calidad.

Las actividades que se realizarán para la supervisión son:

- Verificar la recepción de los materiales de aplicación en las Áreas Estatales de Evaluación y dar seguimiento a la distribución de los mismos al interior de la entidad.
- Verificar las capacitaciones operativas que se efectúen en los estados para la aplicación censal.
- Verificar los procedimientos operativos durante la aplicación.
- Verificar y dar seguimiento a los procedimientos para la recuperación y envío a procesamiento del material utilizado en la aplicación censal.

B. Aplicación de los instrumentos para la Evaluación diagnóstica censal (EDC)¹

1. La SEP enviará a las entidades federativas el material para esta evaluación (las pruebas, hojas de respuestas y manual de aplicación, calificación y análisis de las pruebas).
2. La Subsecretaría de Educación Básica (SEB) informará a los subsecretarios de educación básica de las entidades federativas sobre este programa.
3. Las autoridades educativas estatales coadyuvarán con la distribución de las pruebas a todas las escuelas, a través de la estructura educativa el inicio del ciclo escolar.
4. Las pruebas de Lenguaje y Comunicación y Matemáticas, deberán ser entregadas por los directores escolares a los docentes de 4º grado al inicio del curso, para su aplicación en las fechas que determine la SEB.
5. Las autoridades educativas supervisarán el cumplimiento de esta instrucción a través de la estructura educativa, los supervisores escolares y los asesores técnico pedagógicos (ATP).
6. Los docentes de 4º grado calificarán las pruebas utilizando el software en línea diseñado por la SEP, o con el procedimiento manual establecido para ese propósito, cuyo modelo de puntuación es sencillo y comprensible.
7. Los docentes de 4º grado entregarán los resultados y el diagnóstico del grupo al director de la escuela y éste al Consejo Técnico Escolar.
8. La SEB proporcionará orientación al Consejo Técnico Escolar de cómo proceder a su análisis y al establecimiento de metas y compromisos para los grupos de 3º y 4º grado de educación primaria.

Los resultados obtenidos sólo serán utilizados al interior de cada escuela.

¹ Basado en power point presentado por la DGEPE de la SEP a la Conferencia del Sistema Nacional de Evaluación Educativa el 29 de abril de 2015.

Anexo 2.

Descripción de las escalas de habilidades para la convivencia escolar

La convivencia es una competencia para la vida. Su desarrollo requiere, entre otros elementos empatía, relacionarse armónicamente con otros y la naturaleza; ser asertivo; trabajar de manera colaborativa; tomar acuerdos y negociar con otros; crecer con los demás; y reconocer y valorar la diversidad social, cultural y lingüística.

En México resulta de gran relevancia la evaluación de resultados educativos relativos a la convivencia ya que ha sido incluida en el Plan de Estudios de Educación Básica 2011 como una de las competencias esenciales que deben desarrollar los alumnos a lo largo de la vida, tanto dentro de la escuela como fuera de ella.²

Diversos estudios han señalado que una buena convivencia es un factor de éxito para la adquisición de aprendizajes curriculares. La convivencia escolar constituye un bien en sí mismo, ya que repercute en la formación ciudadana de los estudiantes,³ pues supone la construcción de una experiencia escolar formativa para desarrollar valores, actitudes y habilidades socioemocionales y éticas que sustentan una convivencia social donde todos participan, comparten y se desarrollan plenamente.⁴

Usos apropiados de los resultados de esta evaluación son:

- Informar a los padres, docentes, autoridades educativas locales, estatales y federales y sociedad en su conjunto sobre las habilidades socioafectivas para la convivencia escolar adquiridas por los alumnos.
- Generar estrategias e implementar acciones para mejorar la convivencia en el aula y la escuela.
- Proporcionar insumos para la emisión de directrices por parte del INEE.

² SEP (2011). *Plan de estudios. Educación básica*. México: SEP.

³ Ministerio de Educación de Chile. (2005). *Metodologías de trabajo para el mejoramiento de la calidad de la convivencia escolar*. Santiago, Chile: Ministerio de Educación. Recuperado de http://www.innovemosdoc.cl/desarrollo_institucional/materiales/convivencia_escolar.pdf

⁴ UNESCO (2008). *Convivencia democrática, inclusión y cultura de paz*. Chile: UNESCO.

Usos inadecuados de los resultados de esta evaluación:

- Hacer *rankings* de cualquier tipo.
- Sancionar a las escuelas o a los alumnos.
- Diagnosticar acoso escolar.
- Valorar niveles de convivencia de cada alumno.
- Etiquetar a los alumnos o a las escuelas como violentos.
- Condicionar apoyos económicos o materiales.
- Evaluar a las escuelas, el desempeño docente, programas, u otros componentes del sistema educativo.

A continuación se describen brevemente las escalas que forman parte del Cuestionario del alumno asociado a las pruebas PLANEA aplicadas en 2015 en sexto de primaria y en tercero de secundaria, en las modalidades ELSEN y ELCE. Las escalas fueron seleccionadas y adaptadas a partir de dos fuentes principales: un trabajo previo de validación realizado por Palomar (2015)⁵ a una serie de escalas relacionadas con el tema, y un estudio realizado para el Ministerio de Educación de España.⁶

Competencia social escolar

- La competencia social escolar juega un papel importante en la adaptación de los individuos al entorno escolar influyendo en las relaciones con maestros y pares, así como en el rendimiento académico.
- Los individuos socialmente competentes son aquellos que tienen las habilidades necesarias para resolver problemas en una forma que les permite elegir y comportarse de manera socialmente adecuada.

Conducta prosocial

- La conducta prosocial incluye la consideración social e inclusión de otros a través de actos que implican cooperar, compartir, ayudar, expresar empatía, dar un trato amable y brindar apoyo verbal.
- La conducta prosocial también comprende tomar la iniciativa y defender a otros cuando es necesario.

Relación positiva con pares

- La relación positiva con pares se refiere al apoyo social que se da entre compañeros, lo que se relaciona significativamente con la autoestima y el sentido de pertenencia.
- La autoestima es la percepción valorativa que las personas tienen de sí mismas. Una alta autoestima les permite tomar decisiones responsables sin atender contra su integridad ni la de los demás.
- La pertenencia implica sentirse parte de los distintos grupos en los que se desenvuelve una persona.

⁵ Palomar, J. (2015). *Rendimiento académico, ajuste psicosocial y bienestar subjetivo en adolescentes hijos de beneficiarios del Programa de Desarrollo Humano Oportunidades*, FOSIDES, CONACYT, SEDESOL, PROSPERA. México: Universidad Iberoamericana.

⁶ Díaz-Aguado, Martínez y Martín (2010). *Estudio estatal sobre la convivencia escolar en la educación secundaria obligatoria*. España: Ministerio de Educación.

Prácticas que obstaculizan la convivencia

- Son acciones que impiden la construcción de la paz, tales como las agresiones físicas y verbales y algunas formas de violencia.
- Estas prácticas impiden la solución de conflictos de manera no violenta ya que son acciones en la que se atenta contra la dignidad o bienestar de otros.

Acoso escolar

- Implica la incapacidad de actuar de manera cooperativa y constructiva en contextos escolares. Las víctimas del acoso escolar tienden a tener baja autoestima.
- Destaca tres rasgos: la intencionalidad del acosador de causar daño, el desequilibrio de poder que caracteriza la relación y el carácter repetitivo de la conducta.

Manejo pacífico de conflictos

- Es la forma en que se busca resolver situaciones en las que se presentan diferencias de necesidades, intereses y valores entre dos o más perspectivas que se contraponen.
- Está encaminado a solucionar los conflictos de forma no violenta a través de mecanismos de regulación como el cumplimiento de reglas y el aprendizaje de estrategias tales como el arbitraje, la mediación y la negociación.

Los alumnos que participarán en la ELSN también responderán un conjunto de reactivos que plantean la resolución de dilemas relacionados con la convivencia escolar. El criterio para seleccionar los dilemas fue que reflejaran diferentes posturas frente a situaciones en las que se ponen en juego intereses, necesidades, valores y prácticas que pueden favorecer u obstaculizar el bien común y las relaciones con otros.

Anexo 3.

Estructura de los cuestionarios de contexto

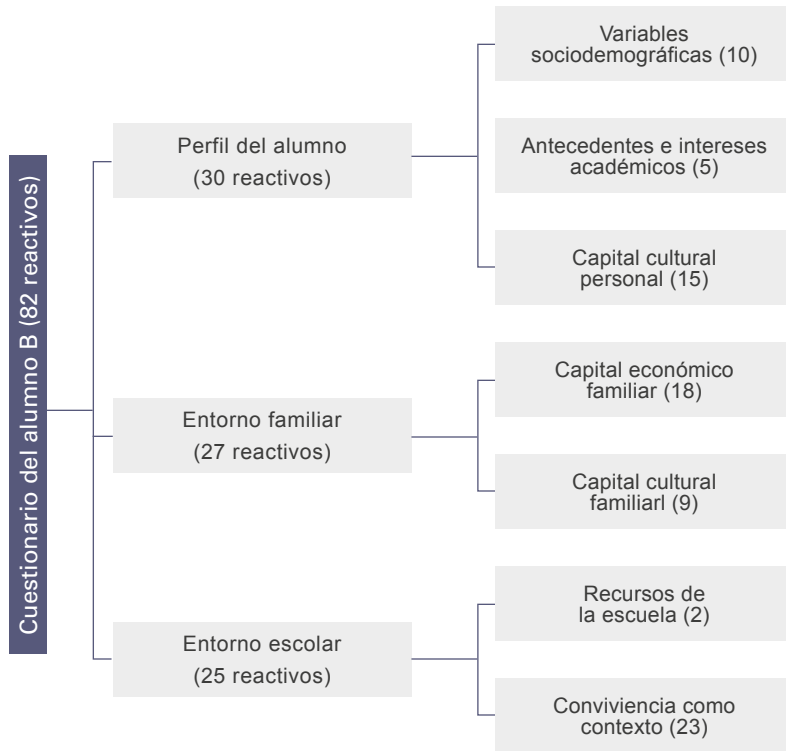
A continuación se presenta la estructura de los cuestionarios de contexto utilizados en 2015 para la aplicación en sexto de primaria y tercero de secundaria en las modalidades ELSEN y ELCE.

Cuestionario del alumno

El cuestionario de contexto del alumno explora tres dimensiones que se relacionan con el logro educativo: el perfil del alumno, con variables sociodemográficas académicas y culturales; el entorno familiar, con variables económicas y culturales; y el entorno escolar, con énfasis en las condiciones para la convivencia en la escuela. A continuación se presentan los diagramas de la estructura general del cuestionario y de cada escala.⁷

⁷ Entre paréntesis se menciona el número de reactivos que conforman cada dimensión y escala.

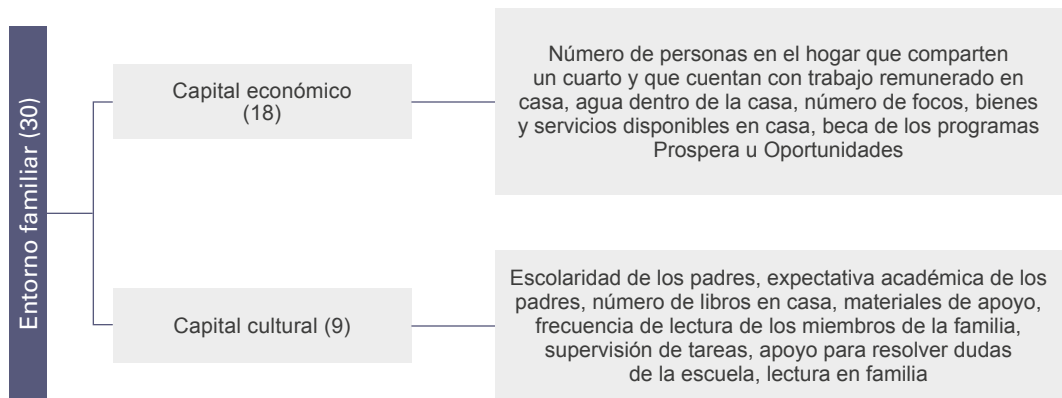
Estructura general del cuestionario de contexto del alumno



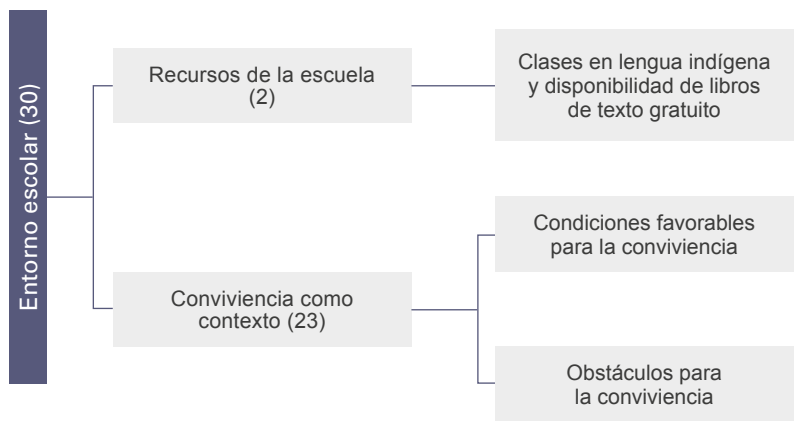
Estructura de la dimensión Perfil del alumno



Estructura de la dimensión Entorno familiar



Estructura de la dimensión Entorno escolar

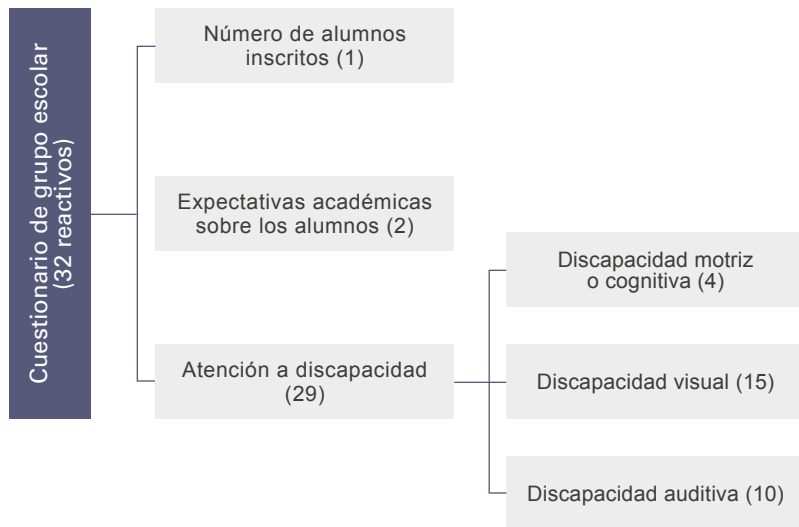


Cuestionario de grupo escolar

Este cuestionario se da a responder a los docentes frente a grupo en el caso de primaria, y a los tutores, coordinadores u otros profesores que conozcan bien al grupo al que se hace referencia en el caso de secundaria. A continuación se presenta un diagrama con la estructura general del cuestionario.⁸

⁸ Entre paréntesis se menciona el número de reactivos que conforman cada ámbito.

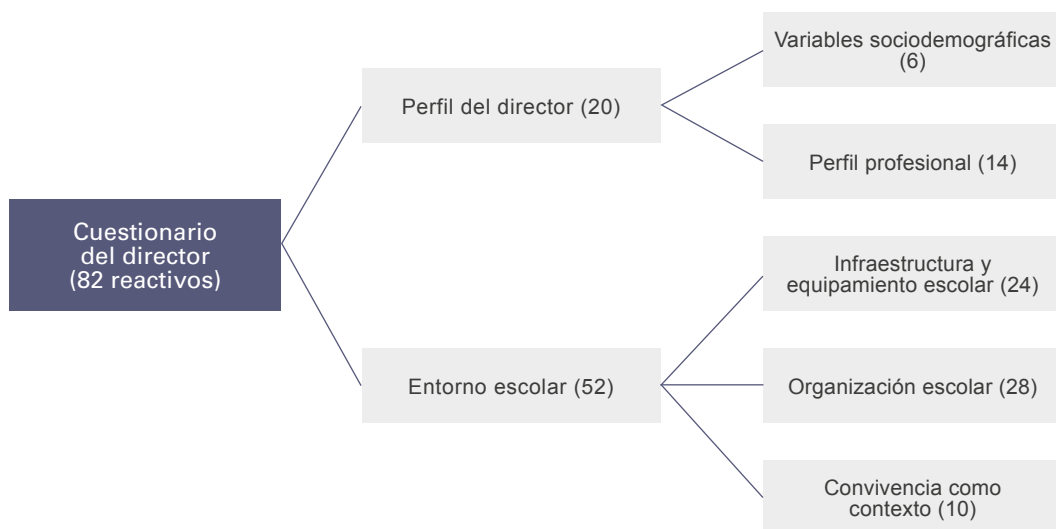
Estructura general del cuestionario de grupo escolar



Cuestionario para directores

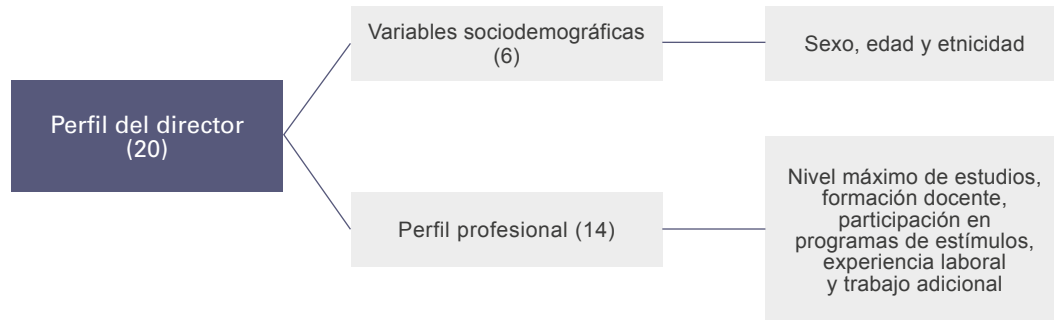
El cuestionario para directores explora dos dimensiones: el perfil del director y el entorno escolar. A continuación se presentan los diagramas con la estructura general del cuestionario y de cada escala.⁹

Estructura general del cuestionario de contexto para directores

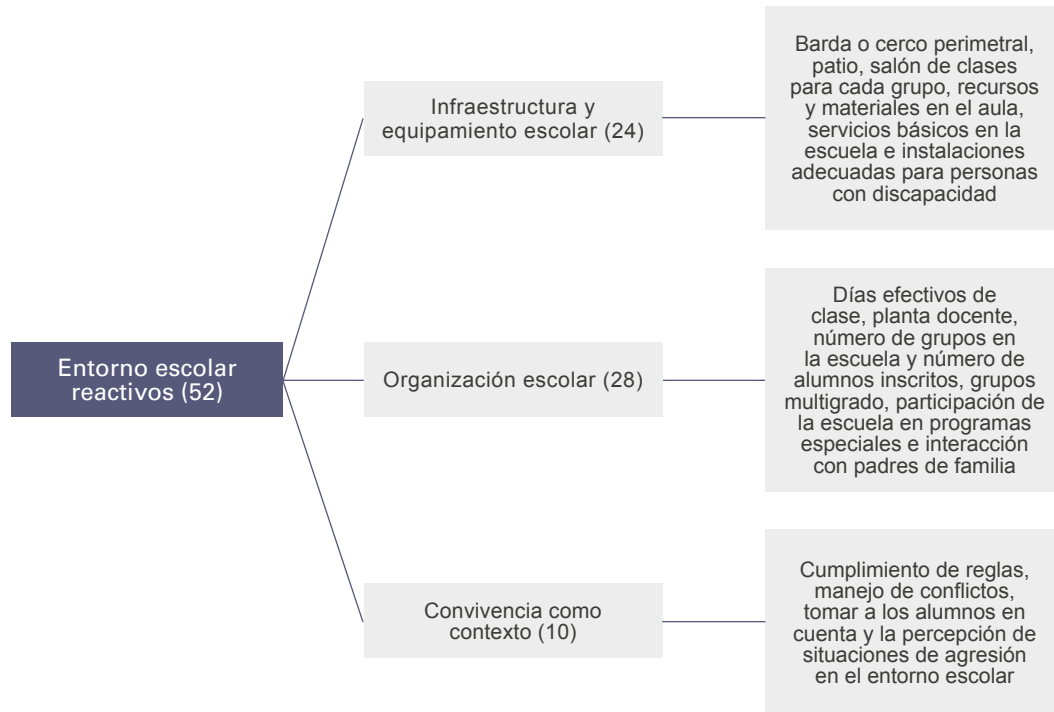


⁹ Entre paréntesis se menciona el número de reactivos que conforman cada dimensión y escala.

Estructura de la dimensión Perfil del director



Estructura general la dimensión Entorno escolar



Anexo 4.

Aspectos centrales del diseño de la muestra para las Evaluaciones referidas al Sistema Educativo Nacional (ELSEN)

Población objetivo

El Instituto Nacional para la Evaluación de la Educación (INEE) define como población objeto de estudio a los alumnos de **sexto grado de primaria** y los alumnos de **tercer grado de secundaria** que estén cursando el ciclo escolar vigente. Para estas subpoblaciones se definen sendas muestras que son independientes entre sí.

El presente Diseño muestral fue elaborado considerando el logro educativo como la principal variable de diseño. El logro educativo, por su naturaleza, no es una variable que se mida directamente por lo que se le da el nombre de 'latente'.

Dominios de estudio

Para la definición de los dominios de estudio, es decir, las subpoblaciones sobre las que se darán resultados, se exploraron variables con las cuales la población de alumnos del país se pudiera desagregar bajo la perspectiva de su potencial para comunicar resultados, y que tuvieran las características técnicas necesarias para reconstruir consistentemente a futuro dichas subpoblaciones. Además, se consultaron los requerimientos y las perspectivas de otras áreas del Instituto con la intención de que dichos dominios se utilicen para otros proyectos y así complementar y contextualizar de mejor manera las evaluaciones que realiza el INEE.

Las variables exploradas para la definición de los dominios de estudio fueron:

- Tipo de servicio, sea para primaria o secundaria
- Sostenimiento de la escuela

- Entidad federativa
- Tamaño de localidad donde se ubica la escuela (INEGI)
- Índice de marginación del Consejo Nacional de Población (CONAPO)
- Escuela (primaria) multigrado: condición de la escuela de tener grupos de primaria —al menos uno— en donde se atiende más de un grado escolar en un mismo grupo (multigrado)

Las variables seleccionadas, debido a sus características para hacer una partición en el marco muestral son:

- Tipo de servicio, sea para primaria o secundaria
- Sostenimiento de la escuela
- Entidad federativa
- Tamaño de localidad (INEGI)

Es importante destacar que el diseño permitirá estimar resultados de los alumnos de escuelas multigrado, y de los alumnos en escuelas en condiciones de marginalidad alta, media y baja. En términos de muestreo, a estos dominios se les conoce como **Dominios no planeados**, en el sentido de que el marco muestral no se segmenta por ellos.

Ejemplo: Es posible dar resultados de logro de hombres y mujeres simplemente porque en la muestra existe una cantidad suficiente de ellos como para hacer una estimación con una precisión aceptable. Para lograr que la muestra permita generar resultados de estos dominios no planeados, se utilizó la estratificación implícita, la cual permite obtener una muestra en la que las unidades muestrales están en las mismas proporciones en las que se encuentran en la población.

A continuación se presentan los resultados de la exploración de las variables utilizadas en la definición de los dominios de estudio.

Tipo de servicio y el sostenimiento

El *tipo de servicio* y el *sostenimiento*, son las clasificaciones que la SEP ha usado tradicionalmente y que se encuentran disponibles en la estadística de la Encuesta de centros escolares que cada año se realiza (Formato 911). Si bien hay otro tipo de clasificaciones como escuelas de tiempo completo, internados, etcétera, éste no está sistemáticamente identificado en la estadística de la SEP. En las siguientes tablas se muestran los totales de alumnos y de escuelas con base en esta clasificación.

Alumnos y escuelas de 6° de primaria por tipo de servicio y sostenimiento

Tipo de servicio	Sostenimiento							
	Alumnos de 6°				Escuelas con alumnos de 6°			
	Público		Privado		Público		Privado	
	Total	%	Total	%	Total	%	Total	%
Comunitarias	15 461	0.62	0	0	7 309	7.84	0	0
Generales	2 144 086	85.74	207 803	8.31	67 983	72.9	8 135	8.72
Educación indígena	1 33 193	5.33	25	0.001	9 820	10.53	2	0.002

Alumnos y escuelas de 3° de secundaria por tipo de servicio y sostenimiento

Tipo de servicio	Sostenimiento							
	Alumnos de 3°				Escuelas con alumnos de 3°			
	Público		Privado		Público		Privado	
	Total	%	Total	%	Total	%	Total	%
Comunitarias	10 111	0.52			2 433	6.59	0	0.00
Generales	831 549	42.72	152 434	7.83	7 145	19.34	4 220	11.42
Migrantes	139	0.01			24	0.06	0	0.00
Técnicas	526 353	27.04	11 034	0.57	4 319	11.69	272	0.74
Telesecundarias	405 173	20.82	611	0.03	18 243	49.39	16	0.04
Trabajadores	8 867	0.46	65	0.00	262	0.71	3	0.01

A partir de los datos de la tabla puede decirse que:

- Las primarias generales concentran a la gran mayoría de alumnos (86%).
- Las primarias comunitarias atienden a menos de 1% de alumnos del país.
- Las primarias privadas son casi en su totalidad generales.
- Las secundarias generales y técnicas concentran a 70% de alumnos y las telesecundarias a 21%.
- Las secundarias comunitarias, de migrantes y de trabajadores atienden respectivamente a menos de 1% de alumnos del país.
- Las telesecundarias representan 50% de las secundarias del país.
- Las secundarias privadas son principalmente generales.

La distribución de alumnos y la de escuelas son diferentes, por ejemplo: las primarias comunitarias concentran menos del 1% de alumnos del país, pero la cantidad de escuelas es casi de 8%.

Esto pone de manifiesto que para efectos del diseño de muestras, se deben tener claros los objetivos de los estudios. En principio, PLANEA deberá aportar información sobre el estado que guardan los aprendizajes de los estudiantes en el SEN, es decir, se debe diseñar una muestra de alumnos con una distribución muy particular. Si en futuras administraciones de PLANEA, se requiere hacer inferencias sobre las características y condiciones de las escuelas del país, con diferente distribución, se deberán hacer los ajustes necesarios.

Debido a que en el levantamiento de PLANEA de 2015 no será incorporado otro estudio, el diseño de la muestra se llevará a cabo con base en la distribución de alumnos.

Entidad federativa

Se tiene contemplado que a partir de la muestra se puedan dar resultados de cada una de las entidades federativas con un nivel de precisión aceptable conforme a los estándares internacionales que más adelante se detallan. Al interior de ellas, y con un nivel de precisión menor, se podrán dar resultados de, a lo más, tres subpoblaciones.

Tamaño de localidad donde se ubica la escuela

El tamaño de la localidad en donde se ubica la escuela se obtuvo con información del INEGI. En la exploración de la variable de tamaño de localidad, se encontró que los alumnos tanto de sexto de primaria como de tercero de secundaria pueden ubicarse en 4 grupos (1 a 499; 500 a 2 499; 2 500 a 99 999 y 100 000 o más habitantes). En las siguientes tablas se muestra su distribución. Una desagregación mayor, entre 2 500 y 99 999 habitantes conduciría a tener categorías vacías en la siguiente tabla. Las primarias y secundarias privadas son generales y se encuentran ubicadas principalmente en localidades de 100 000 o más habitantes.

Alumnos y escuelas de 6° de primaria por tamaño de localidad

Tamaño de localidad	Alumnos de 6°						Escuelas con alumnos de 6°					
	Comunitarias		Generales		Educación indígena		Comunitarias		Generales		Educación indígena	
	Total	%	Total	%	Total	%	Total	%	Total	%	Total	%
No identificable	392	0.02	8 761	0.35	636	0.03	208	0.22	372	0.4	83	0.09
1 a 499	13 736	0.55	162 079	6.48	52 650	2.11	6 613	7.09	22 711	24.36	6 685	7.17
500 a 2 499	777	0.03	333 943	13.35	52 873	2.11	285	0.31	13 869	14.87	2 183	2.34
2 500 a 99 999	363	0.01	748 200	29.92	23 953	0.96	146	0.16	16 771	17.99	788	0.85
100 000 o más	193	0.01	1 098 906	43.95	3 106	0.12	57	0.06	22 395	24.02	83	0.09

Alumnos de 3° de secundaria por tamaño de localidad

Tamaño de localidad	Tipo de servicio											
	Comunitarias		Generales		Migrantes		Técnicas		Telesecundarias		Trabajadores	
	Total	%	Total	%	Total	%	Total	%	Total	%	Total	%
No identificable	100	0.01	5 003	0.26			2 325	0.12	1 170	0.06		
1 a 499	9 270	0.48	8 836	0.45	73	0.004	10 553	0.54	83 167	4.27		
500 a 2 499	534	0.03	41 918	2.15	12	0.001	48 912	2.51	209 380	10.76		
2 500 a 99 999	63	0.003	345 182	17.73	38	0.002	197 317	10.14	81 295	4.18	2 680	0.14
100 000 o más	144	0.01	583 004	29.96	16	0.001	278 280	14.30	30 772	1.58	6 252	0.32

Escuelas de 3° de secundaria por tamaño de localidad

Tamaño de localidad	Tipo de servicio											
	Comunitarias		Generales		Migrantes		Técnicas		Telesecundarias		Trabajadores	
	Total	%	Total	%	Total	%	Total	%	Total	%	Total	%
No identificable	25	0.07	85	0.23	0	0.00	32	0.09	52	0.14	0	0.00
1 a 499	2 313	6.26	241	0.65	13	0.04	167	0.45	7 423	20.10	0	0.00
500 a 2 499	72	0.19	939	2.54	3	0.01	908	2.46	8 782	23.78	0	0.00
2 500 a 99 999	16	0.04	3 797	10.28	5	0.01	1 719	4.65	1 455	3.94	79	0.21
100 000 o más	7	0.02	6 303	17.06	3	0.01	1 765	4.78	547	1.48	186	0.50

A partir de los datos de la tabla previa se puede decir que:

- Los alumnos de sexto grado de primarias generales se reparten en los cuatro grupos (1 a 499; 500 a 2 499; 2 500 a 99 999 y 100 000 o más habitantes).
- Los alumnos de sexto grado de primarias comunitarias se concentran en localidades de 1 a 499 habitantes.
- Los alumnos de sexto grado de primarias de educación indígena se concentran en los grupos 1 a 499 y 500 a 2 499 habitantes, es decir, en un ámbito rural.
- Los alumnos de tercer grado de secundarias generales y técnicas se concentran en los grupos de 2 500 a 99 999 y 100 000 o más habitantes.
- Los alumnos de tercer grado de telesecundaria se concentran en el grupo de 500 a 2 499 habitantes, seguido del de 1 a 499 y de 2 500 a 99 999 habitantes.

A continuación se presentan dos tablas en las que se esquematiza la concentración de alumnos de sexto de primaria y de tercero de secundaria. Las celdas combinadas indican que los alumnos se reparten en proporciones pequeñas y que su unión forma una subpoblación a partir de la cual pueden darse resultados. Lo anterior no implica que en el diseño muestral se vaya a eliminar a los alumnos de las celdas vacías. Esta información sirvió para definir subpoblaciones sobre las cuales se darán resultados.

Concentración de alumnos de 6° de primaria por tamaño de localidad

	Tamaño de localidad			
	1 a 499	500 a 2 499	2 500 a 99 999	100 000 o más
Generales	×	×	×	×
Educación indígena	×			
Comunitarias	×			
Privadas				×

Concentración de alumnos de 3° de secundaria por tamaño de localidad

	Tamaño de localidad			
	1 a 499	500 a 2 499	2 500 a 99 999	100 000 o más
Generales		×	×	×
Técnicas		×	×	×
Telesecundarias	×	×		×
Comunitarias	×			
Privadas				×

A partir de la variable de tamaño de la localidad pueden crearse agrupaciones (subpoblaciones) que sean relativamente estables en el tiempo.

Índice de marginación del Consejo Nacional de Población (CONAPO)

El índice de marginación de CONAPO contempla 5 categorías: muy alto, alto, medio, bajo y muy bajo. Para efectos del diseño muestral, éstas categorías se colapsaron en tres, con la finalidad de tener grupos con suficiente población para estar en posibilidad de decir algo de ellos: alto, medio y bajo.

A continuación se muestra su distribución.

Índice de marginación CONAPO	Primaria				Secundaria			
	Alumnos de 6°		Escuelas		Alumnos de 3°		Escuelas	
	Total	%	Total	%	Total	%	Total	%
No identificable	35 799	1.43	2 248	2.41	40 358	2.07	762	2.06
Alto	981 424	39.25	53 492	57.36	691 352	35.52	20 581	55.72
Medio	757 411	30.29	20 251	21.72	558 494	28.69	7 929	21.47
Bajo	725 934	29.03	17 258	18.51	656 132	33.71	7 665	20.75

En general, el tamaño de la población está asociado al nivel de marginación de las localidades: a menor tamaño, mayor marginación.

La desventaja de utilizar el Índice de marginación como variable de segmentación en el diseño muestral es que la definición del nivel de marginación de las localidades se realiza con base en un análisis factorial de componentes principales de las variables que constituyen el índice. Este procedimiento puede provocar que a futuro, la clasificación de las localidades pueda variar debido a que la desagregación se basa en las varianzas analizadas.

Escuelas multigrado

La definición teórica de escuela multigrado para las primarias generales y de educación indígena a implementar en el presente diseño muestral es la siguiente:

Una escuela es multigrado si al menos en un grupo se imparte más de un grado escolar.

Debido a que en el marco muestral no se incluye la información de las escuelas respecto de si en un grupo o más de éstas se dan clases a más de un grado, se empleó la siguiente definición operacional con la información disponible:

- Se calcula la razón de cantidad de grados escolares que se imparten en la escuela y la cantidad total de docentes de la misma ($\text{Cantidad de grados escolares que se imparten en la escuela} / \text{Cantidad total de docentes}$).
- Dicha razón se categoriza en valores enteros redondeados al entero superior. El uno representa una escuela que no es multigrado, el dos un nivel de multigrado bajo (por ejemplo cinco docentes en la escuela para atender seis grados escolares), etcétera. El seis representa el nivel más severo de multigrado pues un docente atiende seis grados escolares.

Definición de Unidad Primaria de Muestreo

Las unidades primarias de muestreo (UPM) están constituidas por la combinación de la clave del centro de trabajo y el turno.

Dominios de estudio (Definición)

A. Dominios primarios

Para sexto de primaria

- Nacional: Conformado por todos los alumnos de la población objetivo.
- Entidades: Conformadas por la partición que generan las 32 entidades federativas (incluyendo el Distrito Federal).

Para tercero de secundaria

- Nacional: Conformado por todos los alumnos de la población objetivo.
- Entidades: Conformadas por la participación que generan las 32 entidades federativas (incluyendo el Distrito Federal).

B. Dominios secundarios

Conformado por los alumnos que se encuentran en la siguiente clasificación de UPM.

Para sexto grado de primaria

- Cursos Comunitarios (CCO). UPM en primarias administradas por el Consejo Nacional de Fomento Educativo (CONAFE) y atendidas por instructores no necesariamente con formación docente.
- Generales públicas en localidades de 1 a 499 habitantes (GRPT01). UPM en primarias generales públicas ubicadas en localidades de 1 a 499 habitantes.
- Generales públicas en localidades de 500 a 2 499 habitantes (GRPT02). UPM en primarias generales públicas ubicadas en localidades de 500 a 2 499 habitantes.
- Generales públicas en localidades de 2 500 a 99 999 habitantes con alta marginación (GRPT03M0A). UPM en primarias generales públicas ubicadas en localidades de 500 a 2 499 habitantes, cuyo índice de marginación es alto o muy alto.
- Generales públicas en localidades de 2 500 a 99 999 habitantes con media marginación (GRPT03M0M). UPM en primarias generales públicas ubicadas en localidades de 2 500 a 99 999 habitantes, cuyo índice de marginación es medio.
- Generales públicas en localidades de 2 500 a 99 999 habitantes con baja marginación (GRPT03M0B). UPM en primarias generales públicas ubicadas en localidades de 2 500 a 99 999 habitantes, cuyo índice de marginación es bajo o muy bajo.
- Generales públicas en localidades con 100 000 o más habitantes con alta marginación (GRPT04M0A). UPM en primarias generales públicas ubicadas en localidades con 100 000 o más habitantes, cuyo índice de marginación es alto o muy alto.
- Generales públicas en localidades con 100 000 o más habitantes con media marginación (GRPT04M0M). UPM en primarias generales públicas ubicadas en localidades con 100 000 o más habitantes, cuyo índice de marginación es medio.
- Generales públicas en localidades con 100 000 o más habitantes con baja marginación (GRPT04M0B). UPM en primarias generales públicas ubicadas en localidades con 100 000 o más habitantes, cuyo índice de marginación es bajo o muy bajo.
- Educación Indígena No Multigrado (INDM0). UPM en primarias administradas por la Dirección General de Educación Indígena que no se consideran multigrado bajo la definición de la SEP.
- Educación Indígena Multigrado (INDM1). UPM en primarias administradas por la Dirección General de Educación Indígena que se consideran multigrado bajo la definición de la SEP.
- Privadas (PRV). UPM en primarias generales públicas e indígenas de sostenimiento privado.

Para tercer grado de secundaria

- Cursos comunitarios (CCO). UPM en secundarias administradas por el CONAFE y atendidas por instructores no necesariamente con formación docente.
- Generales públicas en localidades de 1 a 2 499 habitantes (GRPT12). UPM en secundarias generales públicas ubicadas en localidades de 1 a 2 499 habitantes.
- Generales públicas en localidades de 2 500 a 99 999 habitantes (GRPT03). UPM en secundarias generales públicas ubicadas en localidades de 2 500 a 99 999 habitantes.
- Generales públicas en localidades con 100 000 o más habitantes con alta y media marginación (GRPT04MAM). UPM en secundarias generales públicas ubicadas en localidades con 100 000 o más habitantes, cuyo índice de marginación es medio, alto o muy alto.
- Generales públicas en localidades con 100 000 o más habitantes con baja marginación (GRPT04MOB). UPM en secundarias generales públicas ubicadas en localidades con 100 000 o más habitantes, cuyo índice de marginación es bajo o muy bajo.
- Técnicas en localidades de 1 a 2 499 habitantes (TECT12). UPM en secundarias técnicas ubicadas en localidades de 1 a 2 499 habitantes.
- Técnicas en localidades de 2 500 a 99 999 habitantes (TECT03). UPM en secundarias técnicas ubicadas en localidades de 2 500 a 99 999 habitantes.
- Técnicas en localidades con 100 000 o más habitantes (TECT04). UPM en secundarias técnicas ubicadas en localidades con 100 000 o más habitantes.
- Telesecundarias en localidades de 1 a 499 habitantes (TELT01). UPM en telesecundarias ubicadas en localidades de 1 a 499 habitantes.
- Telesecundarias en localidades de 500 a 2 499 habitantes (TELT02). UPM en telesecundarias ubicadas en localidades de 500 a 2 499 habitantes.
- Telesecundarias en localidades con 2 500 o más habitantes (TELT34). UPM en telesecundarias ubicadas en localidades con 2 500 o más habitantes.
- Privadas (PRV). UPM en primarias generales públicas e indígenas de sostenimiento privado.

C. Dominios no planeados

- Nacional: Escuelas multigrado y no multigrado.
- Nacional: Escuelas en localidades de alta, media y baja marginación.
- Primarias generales públicas: Escuelas en localidades de alta, media y baja marginación.
- Primarias generales públicas: Escuelas multigrado y no multigrado.
- Primarias de Educación Indígena: Escuelas en localidades de alta, media y baja marginación.
- Secundarias generales: Escuelas en localidades de alta, media y baja marginación.
- Secundarias técnicas: Escuelas en localidades de alta, media y baja marginación.
- Telesecundarias: Escuelas en localidades de alta, media y baja marginación.

Cobertura y exclusiones

Exclusiones de sexto de primaria

- Las UPM que no contaban con algún alumno registrado en sexto grado.
- Las UPM que tenían como valor perdido el total de alumnos en sexto grado.
- Las UPM en escuelas generales públicas en las que no fue posible identificar el tamaño de la localidad.
- Las UPM en escuelas generales públicas en localidades de 2 500 a 99 999 habitantes en las que no fue posible identificar el grado de marginación.
- Las UPM en escuelas generales públicas en localidades con 100 000 o más habitantes en las que no fue posible identificar el grado de marginación.
- Los estratos en los que no se pudo asignar al menos una UPM en muestra.
- Las UPM de la entidad de Oaxaca que no son escuelas comunitarias, debido a que la autoridad educativa estatal manifestó que se reúnen las condiciones necesarias para llevar a cabo el levantamiento de datos.

Exclusiones de tercero de secundaria

- Las UPM que no contaban con algún alumno registrado en tercer grado.
- Las UPM que tenían como valor perdido el total de alumnos en tercer grado.
- Las UPM pertenecientes a escuelas de trabajadores y de migrantes.
- Las UPM en escuelas generales públicas, técnicas públicas y telesecundarias públicas en las que no fue posible identificar el tamaño de la localidad.
- Las UPM en escuelas generales públicas en localidades con 100 000 o más habitantes en las que no fue posible identificar el grado de marginación.

Las UPM de la entidad de Oaxaca que no son escuelas comunitarias debido a que la autoridad educativa estatal manifestó que se reúnen las condiciones necesarias para llevar a cabo el levantamiento de datos.

Precisión de las muestras

La precisión que emplean estudios internacionales en el diseño de sus muestras como *TIMSS (Trends in International Mathematics and Science Study)*, *IEA Civic Education Study* y *PISA* se resume en el siguiente cuadro:

La precisión estándar requerida para dar resultados de cualquier población consta de un tamaño de **muestra efectiva** de 400 alumnos para construir intervalos de confianza del 95% al estimar medias poblacionales, porcentajes y coeficientes de correlación.

- Medias: $M \pm 0.1S$ (donde M es la estimación de la media y S la desviación estándar estimada).
- Porcentajes: $p \pm 5\%$ (donde p es el porcentaje estimado).
- Correlaciones: $r \pm 0.1$ (donde r es la correlación estimada).

Puede revisarse en (IEA, 2004), p.114, (Schulz & Sibberns, 2004), p.45 y (Murphy, Martin; Schulz, Wolfram;, 2006), p.7.

Para el presente diseño muestral de PLANEA, los dominios de estudio que cumplen con esta precisión son los siguientes:

- Nacional
- Dominios primarios
- Dominios secundarios
- Dominios no planeados

Cualquier estimación hecha en otras subpoblaciones diferentes a las mencionadas tendrá que valorarse con base en el error estándar y a la desviación estándar estimadas. Debe considerarse que las subpoblaciones que se obtienen de combinar los dominios arriba citados, tendrán precisiones menores respecto a los estándares previamente estimados, debido a que el tamaño de la muestra disminuye, por lo que se recomienda que las estimaciones que se obtengan se utilicen con cautela.

Debe tenerse en cuenta que el tamaño de la precisión que se ha presentado, únicamente considera al error debido al muestreo. No se toma en cuenta el error de medida de los instrumentos o el error de cualquier otra fuente.

Anexo 5.

Síntesis del análisis de la calidad de los resultados de las Evaluaciones del logro referidas a los centros escolares (ELCE) de PLANEA¹⁰

A diferencia del pasado, cuando los resultados de la prueba ENLACE no se validaron por una institución técnica externa, actualmente el INEE tiene la función de determinar la consistencia y validez de los resultados de la ELCE PLANEA a cargo de la SEP. Como es ampliamente documentado, generalmente razones de costos impiden realizar aplicaciones controladas de pruebas de logro demasiado grandes (en este caso, en censo de escuelas pero no de alumnos) que garanticen que los estudiantes respondan, en forma autónoma e independiente sus exámenes. En esas circunstancias, las pruebas pueden presentar resultados de logro que difieran de forma importante de los que se obtendrían de pruebas controladas (muestrales).

En este documento se presenta una síntesis del análisis estadístico que el INEE implementará para asegurar que los resultados de las evaluaciones de logro referidas a los centros escolares no difieren significativamente de los resultados de las evaluaciones de logro referidas al Sistema Educativo Nacional (ELSEN). Para garantizar la robustez de los resultados, se aplicarán tres análisis estadísticos, dos de ellos confirmatorios del primero. Con esto se podrá asegurar, con altísimo grado de confiabilidad, la existencia de diferencias de importancia en los niveles de logro educativo obtenidos en las dos modalidades de PLANEA.

¹⁰ Este documento es una síntesis de otro más amplio que prepara la Dirección General para la Integración y Análisis de la Información.

Análisis

El primer análisis utiliza como control los resultados de logro educativo obtenidos en la aplicación de la ELSN. Compara si los niveles de logro educativo entre ambas aplicaciones difieren significativamente, en los estratos para los cuales se diseñó la muestra. Por ejemplo, si se reducen en determinada cantidad los puntos porcentuales del primer nivel de logro, denominado aquí I, y aumenta correspondientemente el de algún nivel superior (II, III o IV), siendo además esta modificación estadísticamente significativa a los niveles usuales de confianza (95%).

El segundo análisis que se propone se aplica cuando el primero muestra evidencia de irregularidades, y consiste en la comparación de la distribución de los logros individuales tanto de la prueba que aplicará el INEE como la que aplicará la SEP (*comparación de las distribuciones empíricas*). Este método tiene ventajas respecto al tradicional, basado en una prueba de medias del logro promedio de los exámenes muestrales y censales en los estratos pertinentes. La confirmación de la evidencia resulta en la sospecha de que los estudiantes no contestaron de manera independiente las pruebas.

Además, de la estrategia descrita en el párrafo anterior, se propone un análisis adicional que se aplica cuando el primero confirmó irregularidades, y solo incluye los datos generados por la SEP. A diferencia de los análisis anteriores, este método tiene como insumo la cadena de respuestas a los reactivos de cada uno de los estudiantes, además de los logros, y consiste en formar dos grupos de estudiantes, uno con sospecha y otro sin sospecha de irregularidades, usando el método desarrollado por Wesolovsky (2000) y alternativas desarrolladas por la Dirección General para la Integración y Análisis de la Información del INEE, donde se considera como “irregularidad” la existencia de un *excesivo número de coincidencias en las respuestas de los reactivos entre pares de estudiantes* que sobrepasa a la habilidad del estudiante, la dificultad del reactivo y el azar (aplicado en el aula). Se emplean nuevamente las mismas pruebas de comparación sobre las distribuciones empíricas de los logros entre los grupos formados, con lo que se sustenta aún más la evidencia de irregularidad en la prueba censal.

Análisis adicionales a otros niveles de desagregación, por ejemplo nivel de marginación para cada estado, es tema de investigación que requiere evaluar diferentes aspectos técnicos. Actualmente, se están planeando otras estrategias de análisis para contemplar nuevos niveles de reporte de la prueba.

