

Estándares técnicos para asegurar la calidad en los levantamientos de datos

**Juan Carlos Camacho Gómez
Jorge Sotelo Cortés
María Luz Zarazúa Martínez**

CUADERNO No. 24



**COLECCIÓN CUADERNOS
DE INVESTIGACIÓN**

ISSN 1665-9457

Este texto puede consultarse en: www.inee.edu.mx

MÉXICO, FEBRERO, 2007

CONTENIDO

Introducción	3
Antecedentes	4
El levantamiento de datos como parte del proceso general de la evaluación. Definición y delimitación del proceso	7
Contribución del levantamiento de datos a una medición de calidad	9
◆ Contribución a la confiabilidad	10
◆ Contribución a la equiparabilidad	11
◆ Contribución a la certidumbre	12
Factores que determinan la organización de la aplicación	13
◆ Técnicos. Los objetivos y alcances del estudio y sus respectivos instrumentos	13
◆ Geográficos. Las características geográficas y (demográficas) topográficas del país	14
◆ Políticos. La organización y magnitud del Sistema Educativo Nacional	16
◆ Económicos. Recursos disponibles	17
◆ Éticos. Orientaciones valorales de la organización	18
Evidencias de la calidad de los levantamientos de datos coordinados por el INEE	19
Metodología seguida en la definición de los estándares técnicos para los levantamientos de datos	24
Noción y características de una aplicación de calidad	25
Estándares técnicos para el levantamiento de datos	26
Consideraciones finales	31
Referencias bibliográficas	32

“Los estándares han sido diseñados para guiar el establecimiento del marco institucional, la gestión de la función de evaluación y la realización y el uso de evaluaciones. Representan, asimismo, un referente para las competencias de los profesionales y la ética de trabajo en materia de evaluación, y la intención es que sean aplicados dentro de cada organización”

Estándares de evaluación en el sistema de las Naciones Unidas

1. INTRODUCCIÓN

Las funciones sustantivas del Instituto Nacional para la Evaluación de la Educación (INEE) se organizan en torno a tres núcleos básicos interconectados entre sí: el de indicadores de la calidad educativa, el de evaluación de escuelas y el de evaluación de aprendizajes.

En relación al tercer núcleo, la evaluación de los aprendizajes, y de acuerdo al esquema que se ha ido desarrollando para la delimitación y articulación sistémica de las evaluaciones que realizan los propios planteles escolares, con las que promueven sus autoridades locales o las que desarrollan la Dirección General de Evaluación de Políticas (DGEP) de la Secretaría de Educación Pública (SEP) y el INEE mismo, se ha definido que a este último le corresponde diseñar y aplicar pruebas a gran escala, estandarizadas, en muestras controladas, en ciclos multianuales, con fines de diagnóstico del sistema educativo en su conjunto o una parte del mismo y desde una óptica de evaluador externo.

En este marco, el Instituto realiza estudios de carácter nacional y participa en otros de alcance internacional. En los de tipo internacional, hasta ahora, se ha participado en dos: 1) el Programa Internacional para la Evaluación del Estudiante (PISA por sus siglas en inglés), promovido por la Organización para la Cooperación y Desarrollo Económico (OCDE); y, 2) el Segundo Estudio Regional Comparativo y Explicativo (SERCE), que es una iniciativa encabezada por el Laboratorio Latinoamericano de la Evaluación de la Calidad de la Educación (LLECE) de la UNESCO.

Las evaluaciones nacionales están basadas fundamentalmente en los Exámenes de la Calidad y el Logro Educativo (Excale), desarrollados por el INEE desde 2004 y que se aplicaron por vez primera en mayo de 2005.

Por otra parte, desde su creación el INEE se fijó como meta lograr la más alta calidad técnica en sus trabajos, con prácticas comparables a las seguidas por las instituciones que van a la vanguardia en la materia en el mundo. Lo antes mencionado implica hacer evaluaciones a partir de mediciones válidas y confiables, que se contrasten con referentes válidos y pertinentes, que conduzcan a juicios ponderados y contextualizados.

A su vez, hacer mediciones de calidad requiere cuidar el rigor técnico en todo el proceso, desde el diseño, la construcción y la aplicación hasta la validación de las pruebas, ya que el

relajamiento en la calidad de cualquiera de las fases del proceso tiene efectos perversos en toda la evaluación, hasta llegar incluso a invalidarla.

Se puede tener una evaluación con un diseño robusto, con instrumentos contruidos solidamente, con reactivos de calidad, con una muestra bien diseñada, pero si la aplicación es deficiente y no se logra recabar los datos en las cantidades suficientes requeridas, entonces se afecta toda la evaluación, los resultados no son útiles y los recursos invertidos no logran el producto esperado.

Ahora que el Instituto cuenta con una nueva generación de pruebas, construidas con una serie de lineamientos y estándares rigurosos, los procedimientos de aplicación no pueden quedar al margen, deben también alcanzar altos niveles de calidad y contribuir a que todo el proceso sea sólido técnicamente.

A partir de lo anterior, el propósito de este documento es presentar los estándares técnicos desarrollados para orientar el trabajo de aplicación de los instrumentos de las evaluaciones coordinadas por el INEE. Antes de la presentación de dichos estándares, se hace referencia a diversos aspectos, unos de orden conceptual sobre el proceso de evaluación y otros de orden práctico sobre el trabajo de campo, que deben ser considerados en la definición de procedimientos y estándares para las aplicaciones.

Asimismo, es importante hacer mención de tres consideraciones generales sobre este trabajo:

1. Los estándares que se presentan más adelante son resultado de una corta pero intensa experiencia, ya que a la fecha el INEE ha coordinado 15 operativos de levantamiento de datos, 10 aplicaciones definitivas y 5 piloteos, 10 nacionales y 5 internacionales.
2. Ligado a lo anterior, el desarrollo del INEE, respecto a la aplicación de evaluaciones a gran escala, tuvo como punto de partida la experiencia acumulada por más de 10 años tanto en la DGEP como en las unidades de evaluación de los sistemas educativos estatales.
3. Finalmente, Los estándares no son camisas de fuerza, ni recetas mágicas, ni mucho menos condicionantes del trabajo, constituyen más bien referentes que sirven, por una parte, para apoyar al personal participante y perfilar su trabajo hacia la calidad, y por otra, como parámetro para valorar las aplicaciones desarrolladas por el Instituto.

2. ANTECEDENTES

La evaluación en gran escala que se realiza con el propósito de evaluar los resultados de aprendizaje logrados por los alumnos en los sistemas educativos es relativamente reciente. En algunos lugares, como los Estados Unidos, este tipo de prácticas de evaluación se sistematizó desde la primera mitad del siglo XX, aunque en la mayoría de los países este tipo de evaluaciones comenzaron a cobrar fuerza hasta finales de dicho siglo.

En México desde los inicios de la década de los setentas, la Secretaría de Educación Pública (SEP) comenzó a realizar actividades de evaluación en educación básica, que fueron el antecedente para la creación, en 1972, de una unidad administrativa responsable de esa función y que ha evolucionado hasta llegar a ser la actual Dirección General de Evaluación de Políticas (DGEP). Ese mismo año se aplicaron las primeras pruebas en gran escala para seleccionar a los alumnos de nuevo ingreso a la secundaria. Posteriormente a partir de 1976 se iniciaron las evaluaciones en educación primaria, utilizando muestras nacionales, con el proyecto "Evaluación del rendimiento académico de los alumnos de 4º y 5º grados de educación primaria".

Hasta principios de la década de los 90 la experiencia de la SEP relativa a evaluaciones en gran escala incluía los exámenes de ingreso a la educación normal (IDCIEN), los de Diagnóstico para Alumnos de Nuevo Ingreso a Educación Secundaria (IDANIS) que se aplicaban en sólo tres entidades, y el Estudio Integral de la Evaluación Preescolar, Primaria y Secundaria, aplicado en nueve entidades a 190 docentes.

En 1992 con el establecimiento del estímulo de la carrera magisterial y la evaluación de los factores "aprovechamiento escolar" y "preparación profesional" se hizo necesario aplicar cada año pruebas de conocimientos a números considerables de maestros (más de 500,000 en los primeros años, ahora unos 700,000) y a sus alumnos (cerca de 4,500,000 al principio y poco más de ocho millones actualmente). Según lo señala la propia DGEP¹, en la memoria elaborada con motivo de su aniversario número 30, la instauración del Sistema de Evaluación del Programa de Carrera Magisterial trajo consigo un impulso decisivo para apoyar la consolidación de la evaluación de la educación básica.

En noviembre de 1996 se llevó a cabo el primer levantamiento del proyecto Evaluación de la Educación Primaria (EVEP), con la participación de 26 entidades. Se aplicaron pruebas de español y matemáticas a alumnos que habían concluido 2º y 5º grados en julio anterior. Asimismo, se administraron instrumentos de contexto a padres, docentes y directores, y se recabaron cédulas descriptivas de las escuelas con el fin de relacionar los resultados de la evaluación con las condiciones de los planteles.

Entre 1996 y 2000 se completó el resto de las asignaturas académicas, se cubrieron todos los grados y se alcanzó la cobertura nacional. Anualmente se aplicaron alrededor de 110 000 pruebas en aproximadamente 3 100 escuelas.

Otro hecho que contribuyó de manera importante al avance de las actividades de evaluación, es la determinación de la SEP de impulsar, desde 1996, la creación o consolidación de Áreas Estatales de Evaluación (AEE) en todas las entidades del país, para incrementar la capacidad de utilizar los resultados de las evaluaciones hechas por la DGEP y de hacer otras propias. También promovió la constitución de un mecanismo de colaboración entre los responsables de dichas AEE, el llamado Comité Técnico Interestatal de Evaluación, CTIE.

¹ Antes llamada Dirección General de Evaluación.

Con el estímulo que representó el ingreso de México a la OCDE en 1995, el país comenzó a participar en proyectos internacionales de evaluación educativa: ese mismo año se participó en el Tercer Estudio Internacional de Matemáticas y Ciencias (TIMSS por sus siglas en inglés) coordinado por la Asociación Internacional para la Evaluación del Logro Educativo (IEA), que abarcó a 40 países; en 1997, junto con otros 12 países, en el Laboratorio Latinoamericano de Evaluación de la Calidad Educativa de la OREALC; y desde el 2000 en el proyecto PISA, de la OCDE. El elevado nivel técnico de las evaluaciones internacionales representa a la vez un reto importante y una excelente oportunidad de aprendizaje.

Otro gran paso fue el desarrollo de un estudio que pretendía medir el nivel de desarrollo de habilidades logradas por los alumnos en comprensión lectora y la resolución de problemas matemáticos, con base en el modelo llamado evaluación con referencia a criterio. A partir del ciclo escolar 97-98 en primaria y del 99-2000 en secundaria se inició la aplicación de las Pruebas de Estándares Nacionales de Lectura y Matemáticas, que se acompañaron de instrumentos para obtener información sobre el contexto de la enseñanza, con propósitos explicativos. En el 2001 el volumen de aplicación fue de más de 210 000 alumnos en poco más de 4 500 planteles.

Un salto cualitativo importante fue la creación del INEE. Las Pruebas de Estándares Nacionales coordinadas por la SEP pasaron a ser responsabilidad del Instituto, en el 2004 se hizo cargo directamente de su aplicación en muestras de escuelas y alumnos de 6° de primaria y 3° de secundaria, así como a una muestra de escuelas PEC. Estas aplicaciones fueron las últimas en que el Instituto utilizó los instrumentos de estándares nacionales. En 2005 el INEE comenzó la aplicación de instrumentos nuevos, desarrollados a partir de 2004, los llamados Exámenes de la calidad y el Logro Educativo (Excale) y cuyo principal avance fue tener la posibilidad de evaluar en extenso el programa curricular de estudios.

En el ciclo 2005-2006 la SEP inició el proyecto de Evaluación Diagnóstica Censal, que se propone la aplicación masiva, por parte de los mismos docentes y al inicio del año escolar, de instrumentos de evaluación que permitan al maestro conocer la situación en que recibe a sus alumnos, para que pueda planificar su trabajo adecuadamente.

Al final de este mismo ciclo se aplicaron las pruebas de la Evaluación Nacional del Logro Académico en Centros Escolares (ENLACE), que responde a la demanda social de contar con resultados comparables a nivel de alumno y de plantel educativo, a la vez que ser un insumo importante para apoyar la mejora de los resultados educativos desde la escuela.

A manera de conclusiones, de este repaso rápido del desarrollo de la evaluación de la educación básica en México, y particularmente respecto a sus implicaciones para el levantamiento de datos, es posible afirmar que:

- ◆ Con el aumento progresivo en el volumen de las aplicaciones se incrementó también la complejidad de la tarea, situación que se agravó por la escasez de especialistas y la falta de experiencia de los responsables, que debieron aprender sobre la marcha. Influyeron

también otros factores, burocráticos y administrativos, que afectaron las condiciones de trabajo e interfirieron en ocasiones con los delicados procesos de aplicación.

- ◆ Además de la cantidad, la calidad de la evaluación también presentó un avance importante ya que el diseño de las muestras permitió hacer análisis por entidad federativa y según varios estratos o tipos de escuela, como públicas y privadas, urbanas y rurales o escuelas indígenas, aunque en esa misma medida se generaron nuevas exigencias para el proceso de aplicación.
- ◆ Respecto a la periodicidad, la aplicación anual en varios o todos los grados es pesada, costosa y no indispensable, porque los cambios en las escuelas no ocurren con tanta rapidez. Diseños adecuados permiten análisis robustos con menos aplicaciones.
- ◆ Los procesos de aplicación han sido cuidadosos, pero problemas administrativos han provocado que las pruebas se hayan aplicado en unos años al fin del ciclo escolar y en otros al inicio del mismo. Junto con otros factores, lo anterior imposibilita las comparaciones a lo largo del tiempo.
- ◆ La iniciativa de nuestro país de participar en evaluaciones internacionales, fue un paso en la dirección de fortalecer la evaluación. Dado que esas evaluaciones son hechas por grupos que reúnen a algunos de los más destacados especialistas, esta puede ser una vía muy eficiente de capacitación y aprovechamiento de la experiencia acumulada sobre el tema.
- ◆ El diseño y aplicación de las pruebas Excale y ENLACE, constituyen una nueva era en la evolución del desarrollo de la evaluación en el Sistema Educativo Mexicano, se complementan, siguen rigurosos estándares de calidad y tienen como premisas la difusión y uso de resultados para la rendición de cuentas y la mejora de la calidad. Por otra parte son un paso importante para reducir la diversidad y traslape, en tiempo y propósitos, de las evaluaciones que se realizan actualmente.
- ◆ No hay evidencias suficientes de la calidad de las aplicaciones por lo que el reto es sistematizar y documentar los procedimientos seguidos para dar certeza de que la evaluación está libre de sesgos atribuibles a las actividades relacionadas con el levantamiento de datos.

3. EL LEVANTAMIENTO DE DATOS COMO PARTE DEL PROCESO GENERAL DE LA EVALUACIÓN. DEFINICIÓN Y DELIMITACIÓN DEL PROCESO

No es trivial señalar que la recolección de datos tiene sentido como parte del proceso general de evaluación. En términos prácticos lo anterior significa que los lineamientos y procedimientos de la aplicación deben considerar los propósitos generales de la evaluación, la población objetivo, las características de los instrumentos y las especificaciones para ser

administrados. Consecuentemente, las orientaciones para el trabajo de campo son diferentes según el estudio de que se trate, en otras palabras, la instrucción que puede ser válida para una aplicación no necesariamente lo es para otra.

Suele ocurrir, con más frecuencia de lo deseable, que la estructura responsable de realizar el operativo o no conoce las características de la evaluación o las omite al momento de hacer su trabajo y actúa por extrapolación de situaciones, es decir, como se hizo la vez anterior o en tal otra aplicación, sin que necesariamente sea adecuado.

De acuerdo a la metodología adoptada por el INEE para la construcción de los Excale (Backhoff y colaboradores 2005), el proceso general comprende siete fases, que a su vez se desagregan en 16 etapas, tal como se ilustra en la tabla 1.

Tabla 1. Proceso de diseño, construcción, aplicación y validación de los Excale

Fases	Etapas
I. Planeación general	1. Diseño del Plan General de Evaluación 2. Diseño y elaboración de cuestionarios de contexto 3. Diseño y desarrollo del sistema informático
II. Estructuración	4. Diseño de las pruebas 5. Especificaciones de reactivos
III. Construcción de reactivos	6. Elaboración de reactivos 7. Validación de reactivo 8. Piloteo de reactivos y cuestionarios de contexto
IV. Conformación	9. Selección de reactivos e integración de bloques y formas 10. Edición, armado e impresión
V. Aplicación	11. Muestreo poblacional 12. Administración de las pruebas y captura de resultados
VI. Análisis e interpretación de los resultados	13. Análisis iniciales de resultados 14. Establecimiento de niveles de logro 15. Elaboración de informe técnico de resultados de aprendizaje
VII. Validación	16. Estudios de validez de procesos y resultados de las pruebas

Desde esta propuesta, los productos de una etapa son el insumo para la siguiente, así pues la fase V se refiere a la aplicación que inicia cuando los instrumentos han quedado debidamente integrados y concluye cuando se genera un archivo electrónico que contiene los datos sistematizados y depurados de las respuestas aportadas por los sujetos de la evaluación.

La fase de aplicación se desarrolla en dos etapas, el muestreo poblacional y la administración de la prueba y captura de resultados, que a su vez cada una de ellas se

desagrega en varias actividades, en las cuales intervienen una cantidad considerable de personas, muchas de ellas reclutadas únicamente para el trabajo de administración de las pruebas, pues las evaluaciones externas de gran escala difícilmente pueden ser soportadas por la estructura de la instancia evaluadora, dada la cantidad de personal que requieren.

La etapa que resaltamos, especificamos y describimos en este texto es precisamente la de *aplicación o levantamiento de datos*, que se refiere a las actividades que hay que realizar y pormenores a considerar para poner un instrumento de evaluación o diagnóstico, previamente diseñado por grupos de especialistas, en un aula o centro escolar, para que sea respondido por el alumno, su maestro, el director o padre de familia, según sea el caso; y luego, su respectivo regreso hasta el laboratorio de lectura óptica o procesamiento de datos. Lo anterior, buscando siempre que el instrumento se aplique en las mismas condiciones y tiempos en todos los lugares y se salvaguarde desde luego su seguridad y confidencialidad, destacando por ello la importancia de que al enviar 500 o 90,000 instrumentos de un estudio a campo, no se pierda uno solo.

La diversidad de actividades, los múltiples factores que influyen en cada una de ellas, el número de persona que intervienen en el proceso, así como los aspectos que hay que cuidar de datos, cantidades, tiempos y procedimientos, hace necesario manejar una serie de estándares para garantizar que los resultados en general estén libres de sesgo por factores atribuibles al levantamiento mismo y que cada procedimiento emprendido contribuya efectivamente a dos de los factores que auspician una buena evaluación, la confiabilidad y la equiparabilidad.

4. CONTRIBUCIÓN DEL LEVANTAMIENTO DE DATOS A UNA MEDICIÓN DE CALIDAD

Como se ha mencionado, el proceso de levantamiento de datos debe guardar cierta consonancia con las características de la evaluación, dicha correspondencia se da en dos planos, uno general y otro específico. En el plano general, la logística del levantamiento de datos debe considerar la contribución del proceso a las características generales de toda medición, esto es la confiabilidad, la equiparabilidad y la certidumbre; los siguientes incisos de este apartado describen con mayor detalle cada uno de estos elementos.

En el otro plano, en lo específico, se debe considerar el tipo de información a recolectar, las características de los instrumentos a usar, las fuentes de información y el contexto en que estas se encuentran, esto se trata con mayor detalle en el apartado 5 de este documento.

4.1 Contribución a la confiabilidad

Una buena prueba debe tener dos características para que se le pueda considerar como buen instrumento de medición: validez y confiabilidad.

La confiabilidad es el grado en el que la aplicación repetida de un instrumento de medición al mismo fenómeno genera resultados similares. Se puede hacer una analogía con un termómetro, si por ejemplo medimos cuatro veces la temperatura en periodos de dos minutos y los resultados son muy variables (22°, 10°, 40° y 25° C), es evidente que la medición no es confiable, debido a que el instrumento no sirve o porque se está usando mal.

En términos generales la confiabilidad está relacionada con la consistencia en las mediciones, es decir un instrumento de medición es confiable si cada vez que se utiliza arroja resultados similares. Por supuesto que la consistencia implica que se trate de la misma prueba o alguna forma equivalente, del mismo individuo evaluado u otro comparable, en un mismo momento o en otro diferente.

La confiabilidad de un instrumento de medición es consecuencia de la construcción del instrumento mismo, pero también de la forma como se administra y de la forma en que se leen sus resultados.

En otras palabras, se considera que una medición tienen error si produce inconsistencia en las calificaciones de una situación a otra y puede haber errores inherentes a la prueba, errores asociados a las condiciones de aplicación de la prueba y errores debidos a fluctuaciones en las características del individuo.

A continuación se enuncian algunos aspectos que deben observarse en el diseño de los procedimientos para evitar sesgos en la medición:

Tabla 2. Elementos que apoyan la confiabilidad de la medición

Del instrumento	<ul style="list-style-type: none"> ◆ Equivalencia de las formas en contenido, formato, longitud, dificultad y variabilidad ◆ Claridad de las instrucciones ◆ Tamaño, color y tipo de letra ◆ Tipo de respuestas, forma y lugar de registro
Del procedimiento	<ul style="list-style-type: none"> ◆ Manuales con procedimientos claros ◆ Selección de alumnos ◆ Tamaño del grupo y número de aplicadores ◆ Registro del tiempo ◆ Uso de material de apoyo (calculadora) ◆ Descansos durante la aplicación
Del aplicador	<ul style="list-style-type: none"> ◆ Respeto de los tiempos ◆ Apego estricto a los procedimientos e instrucciones ◆ Asesoramiento que se da al aplicar las pruebas puede entenderse mal y provocar errores en la medición. ◆ Control del material
Del evaluado	<ul style="list-style-type: none"> ◆ Comprensión de las instrucciones ◆ Enfermedad, fatiga. ◆ Ansiedad debida a las pruebas. ◆ Experiencia en la solución de instrumentos del tipo ◆ El entrenamiento previo para registrar las respuestas ◆ La motivación del examinado para dar su mayor esfuerzo. ◆ Intento de copiar ◆ Ausencia parcial ◆ Atención a la diversidad (capacidades especiales interculturalidad)
Del ambiente	<ul style="list-style-type: none"> ◆ Las condiciones del lugar ◆ Interrupciones inesperadas ◆ Condiciones de aplicación. La temperatura, el ruido, ◆ Información expuesta en el salón

4.2 Contribución a la equiparabilidad

La exigencia de los tomadores de decisiones y de la sociedad en general hacia las personas o instituciones que realizan evaluación educativa, es por información sobre la calidad de la educación que determinado sistema educativo, o parte de él, está logrando, así como acerca de su evolución en el tiempo.

Relacionado con esto último, para saber si los niveles de aprendizajes logrados por los alumnos han mejorado o retrocedido, se requiere contar con instrumentos de medición comparables.

La comparabilidad entre una evaluación y otra se logra usando técnicas de equiparación que consisten en utilizar diseños de investigación y métodos estadísticos desarrollados para tal fin, pero también implican considerar algunos otros requisitos observables en la fase de aplicación de los instrumentos.

Martínez Rizo (2004) propone un listado de 17 criterios que se deben cuidar para lograr una buena equiparación, de estos, a continuación se señalan los que tienen relación con la aplicación de los instrumentos de medición:

- ◆ Los ítems a utilizar se pilotearon
- ◆ Las características formales de las pruebas (tipografía, formato, etcétera) son adecuadas y se mantienen iguales en las aplicaciones sucesivas.
- ◆ Hay procedimientos estándar de aplicación y medidas de seguridad, las cuales evitan que algunos sustentantes conozcan previamente la prueba o algunos ítems, asegurando la igualdad de condiciones de los sustentantes y la ausencia de sesgos. Los procedimientos se mantienen idénticos a lo largo del tiempo.
- ◆ La aplicación de las formas tiene lugar en momentos similares del año.

4.3 Contribución a la certidumbre del estudio

El levantamiento de datos es la parte más visible del proceso de evaluación, ya que las fases previas y posteriores a ésta, son desarrolladas por el equipo responsable del estudio y en ciertas actividades se extiende solo a reducidos grupos de expertos, jueces, analistas, informantes o proveedores.

En el levantamiento de datos se incorporan a la organización y trabajo coordinadores estatales y regionales, supervisores y aplicadores; además, participan, o al menos se enteran, autoridades educativas, supervisores, directores, profesores, alumnos y padres de familia, que sumados todos llegan a ser varias decenas de miles, según el tamaño de la muestra. Cabe señalar que aunque no es lo deseable, en no pocos casos este es el único momento en que la comunidad escolar tiene conocimiento de la evaluación.

La organización del levantamiento de datos debe hacerse ofreciendo información clara y suficiente a los diversos actores involucrados, dando muestras del rigor técnico y cuidando aspectos que hagan evidente la formalidad del proceso.

Dada la incipiente cultura de evaluación que aún persiste en nuestro país, es fácil que un supervisor, un director, un profesor o algún padre de familia descalifique los resultados de toda la evaluación, y seguramente influirá en otros actores a su alrededor, solo porque el aplicador llegó tarde, porque la caja contenía un cuadernillo menos, porque no fueron evidentes las medidas de seguridad del material, porque un cuestionario tenía fallas de impresión, porque se enteraron de la aplicación hasta el momento en que llegó el aplicador, porque los datos de identificación del plantel no coincidían, o por cualquier otro error.

Por lo anterior, durante la aplicación se deben reducir al máximo las irregularidades a causa del material, la organización o los procedimientos, en cambio se debe lograr que todo el proceso brinde certidumbre.

5. FACTORES QUE DETERMINAN LA ORGANIZACIÓN DE LA APLICACIÓN

El proceso de levantamiento de datos para la evaluación supone poner en movimiento una buena cantidad de recursos materiales, humanos y financieros de manera planificada a fin de lograr que todo esté dispuesto para que la aplicación ocurra en tiempo y forma, y que el material de evaluación se recolecte integra y oportunamente.

El proceso en sí mismo es complejo, pero se hace aún más debido a los siguientes factores:

5.1 *Técnicos.* Los objetivos y alcances del estudio y sus respectivos instrumentos.

Al definir los alcances y objetivos de los estudios de evaluación, se determina también de manera indirecta, el conjunto de elementos normativos que deberán seguirse para organizar el levantamiento de datos.

En general podemos decir que los estudios de evaluación que lleva a cabo el INEE son muestrales, de gran escala y que presuponen una aplicación de sus instrumentos en forma estandarizada, pero una gama de variaciones aparecen al respecto cuando se definen ciertas características esenciales del estudio.

Un primer caso se presenta al identificar si el tipo de aplicación es un piloteo o un ejercicio definitivo, en cuyo caso las observaciones sobre los tiempos y formas de administrar los instrumentos es evidentemente distinto.

Otra variante de las aplicaciones se deriva de la población objetivo del estudio, donde en frecuentes ocasiones el actor principal es el alumno, pero en otras puede ser el docente o el director del plantel o una combinación de ambos, además de los diferentes niveles educativos, modalidades y grados escolares de estudio. Esto debe ser previsto también en la organización del operativo para la obtención de datos.

Una tercer variante se puede identificar con el impacto mismo al que atiende el ejercicio de evaluación y de lo cual se desprende también el nivel de seguridad que deben tener los materiales de evaluación. En este sentido resulta de vital importancia conocer si el estudio es de alto impacto y si de sus resultados dependerá por ejemplo: el ingreso a cierto nivel de estudios, la obtención de un puesto laboral o la promoción a un mejor nivel de ingresos. En correspondencia a lo anterior se puede decir que ello requerirá un alto nivel de seguridad sobre el manejo de sus materiales de evaluación (Gajardo, 2002). Por

el contrario, se pueden tener evaluaciones de bajo impacto para el evaluado, con niveles de seguridad menos estrictos aunque no por ello, sin observancia de ciertos estándares mínimos.

Algunas otras situaciones técnicas que de alguna manera influyen en el proceso de aplicación son referentes a tener instrumentos guiados o autoadministrables; de conformación matricial con varias versiones o de un solo tipo; de respuesta abierta u opción múltiple; de aplicación censal por grado o grupo o con selección de muestras a su interior; con o sin reemplazo de escuelas o alumnos. Las implicaciones de lo anterior se transforman en supuestos y posteriormente en normas de ejecución durante los levantamientos de datos.

Dichos elementos normativos se establecen y agrupan como lineamientos a seguir durante el proceso del levantamiento y se pueden clasificar en cuatro tipos como se señala a continuación:

1. Lineamientos sobre los instrumentos.- se refiere a aquellos que dan cuenta de la calidad de los materiales de evaluación desde su fase de impresión, empaquetado y distribución, hasta la recuperación de lo mismos una vez utilizados y su posterior entrega al área de lectura e interpretación. Sobre este particular se puede resaltar que los propósitos de evaluación marcan el tipo de instrumentos a utilizar y por ende, el nivel de cuidado de los mismos, así como su debida confidencialidad.
2. Lineamientos de operación.- aquellos que son utilizados en el momento de la aplicación, que implica: la administración de los instrumentos, lectura de instrucciones, tiempos de resolución, orden en que se entregan y recogen, además de los casos o situaciones especiales donde no se hace entrega de los mismos a la población objetivo entre otros.
3. Lineamientos de registro y control.- Se refiere a las especificaciones sobre el uso y manejo de registros del proceso, mismos que constituyen la evidencia documentada de la calidad de proceso y son fuente de consulta para explicar anomalías detectadas en la lectura.
4. Lineamientos de restricción.- los que deben obedecerse a fin de no caer en situaciones de riesgo que provoquen el sesgo de la información derivada del proceso de aplicación, tales como no llegar tarde a los centros de aplicación por parte de los Aplicadores, no dejar materiales sueltos en la escuela, no fotocopiar ni permitir lo mismo sobre el material de evaluación utilizado en las escuelas, no dar consejos de resolución a los alumnos, etc.

5.2 Geográficos. *Las características geográficas y topográficas del país.*

El territorio nacional tiene cerca de 2 millones de Km², en donde se presentan casi todos los paisajes naturales que son posibles de encontrar en el planeta. Desde las condiciones

más áridas representadas en las zonas desérticas del norte, hasta las selvas más exuberantes de las zonas húmedas, desde los matorrales tropicales más cálidos hasta los paramos de altura como sierras y montañas. Asimismo, el país está dividido políticamente en 32 entidades, 2402 municipios y alrededor de 199 391 localidades.

Las consideraciones del hábitat se convierten en un verdadero punto de atención cuando se quiere asegurar que las condiciones de aplicación sean las mismas o lo más parecidas posibles en los centros escolares y ello refiere necesariamente a la logística de la aplicación. Algunas de estas consideraciones que deben siempre tomarse en cuenta son:

- ◆ El país cuenta de manera oficial con 3 usos horarios, el del centro que es el que abarca la mayor parte del País, el de las montañas, con una hora de retraso con respecto al primero y el horario del pacífico, con dos horas de retraso respecto al horario del centro. Esto implica tener siempre en consideración que la administración de los instrumentos variará en su hora de aplicación según la entidad y no será el mismo tiempo si se trata de Aguascalientes que de Nayarit o Baja California.
- ◆ Las extensiones de las entidades federativas son muy diversas, por lo que el tiempo de traslado para llegar a algunas comunidades requiere de parte de ciertos Aplicadores, el trasladarse desde uno o dos días antes, en tanto que otros solo invertirán el tiempo de cruzar algunas calles.
- ◆ Así mismo, estas diversas distancias de recorrido deben considerarse para anticipar el tiempo necesario de envío de los materiales de aplicación a las entidades, pues aparte de la entrega de estos en la capital por parte del INEE, existe un recorrido incluso en ocasiones mayor, que se hace de la capital de aquellas a cada una de sus propias regiones y de ahí a las escuelas. Tales son los casos de Veracruz, Coahuila y Chihuahua y algunas regiones específicas de Chiapas, Michoacán, Oaxaca o Guerrero.
- ◆ Es menester resaltar también que algunas rutas en centros urbanos grandes como Guadalajara, Monterrey y el Distrito Federal pueden representar para los Aplicadores traslados que les ocupan hasta cerca de tres horas, lo cual es relevante si su presencia en las escuelas es necesaria unos 15 minutos antes de la hora de entrada, es decir a las 7:45.
- ◆ La temperatura en algunas entidades alcanza niveles muy altos, por lo que las clases empiezan una hora antes de lo normal, sobre todo al hablar de las escuelas primarias, como es el caso de Tabasco y Veracruz. Por el contrario en ciertas regiones de montaña hacen que el inicio de clases se retrase hasta en una hora, como sucede en Chihuahua, Zacatecas y algunas regiones de Michoacán o Durango.
- ◆ Los montos de recursos requeridos para la aplicación en una sola escuela con condiciones de difícil acceso, pueden ser iguales a los de otras 10 escuelas juntas de carácter urbano o incluso más, como es el caso de Durango, donde el traslado a ciertas escuelas desde la capital del estado implica para un Aplicador la renta de una avioneta y un guía para llegar a su destino.

La combinación de las anteriores situaciones debe tomarse en cuenta a la hora de definir normas específicas de organización, respecto de tiempos de aplicación y previsiones presupuestales entre otras.

5.3 Políticos. *La organización y magnitud del Sistema Educativo Nacional.*

En México existen 208 830 escuelas de educación básica, a las que asisten cerca de 24.9 millones de alumnos de entre 3 y 16 años de edad, atendidos por poco más de un millón de docentes.

Por otra parte, en el subsistema de educación básica coexisten por razones históricas diferentes tipos de servicio, por lo que funcionan planteles públicos y privados, urbanos y rurales, matutinos y vespertinos, específicos para atender población indígena y otros apoyados en la tecnología de la televisión, escuelas de organización completa, incompleta y otros de atención multigrado, todos ellos con un programa educativo de carácter nacional.

Sin embargo, es importante señalar que la administración de la educación básica tiene un carácter descentralizado y que las escuelas públicas, que representan el 89% del total de escuelas (SEP 2006), dependen directamente de los gobiernos estatales.

La situación en educación media superior es más compleja, ya que no sólo hay mayor diversidad de modalidades, sino también diferentes planes y programas de estudio, tiempo de duración e instancias de coordinación.

Lo anterior tiene al menos tres implicaciones para el proceso de levantamiento de datos, una respecto a los estratos de la muestra según la representatividad que se quiera por tipo de escuela; una segunda respecto a las instancias con quienes se deben coordinar las acciones; y otra más respecto a ciertos procedimientos específicos en función del nivel y modalidad de servicio educativo de que se trate.

Por otra parte, dadas las características de la estructura organizativa del INEE, la coordinación del trabajo de campo para el levantamiento de datos descansa primordialmente en las áreas de evaluación de las entidades federativas, que nombran un Coordinador Estatal del estudio y quien apoyado a su vez en Coordinadores Regionales recluta, capacita y coordina a las personas que administran finalmente los instrumentos de evaluación en las escuelas.

Como es imaginable, estas circunstancias demandan la coordinación de diversas instancias, la participación de muchas personas, el asegurar que los contenidos de información sean claros para todos los involucrados y la superación de varios acuerdos de distinto nivel para asegurar que los levantamientos arrojen datos confiables sobre la calidad de la educación en México.

Algunas de las grandes acciones que buscan alcanzar esta uniformidad y el equilibrio necesario entre las diversas dimensiones sociales y políticas existentes son:

- ◆ El acercamiento constante entre la Dirección de Relaciones Nacionales del INEE y las Áreas Estatales de Evaluación dependientes de los sistemas educativos locales para retomar de ellas las inquietudes principales sobre las evaluaciones
- ◆ La invitación formal a cada entidad federativa para participar en cada estudio
- ◆ La transmisión uniforme de la información a través de sesiones estructuradas y monitoreadas de capacitación, folletos informativos y comunicados externos oportunos con las personas que fungen como Coordinadores Estatales, responsables a su vez por lo general de las Áreas Estatales de Evaluación.
- ◆ Información amplia y precisa entre las áreas internas del Instituto para alcanzar la efectividad de las acciones y responder oportunamente los requerimientos de los estados.
- ◆ El diseño, entrega y uso de manuales sobre el desarrollo de los procedimientos para las distintas figuras centrales que apoyan en el levantamiento de datos y la aplicación misma.
- ◆ El envío de recursos económicos a los Coordinadores Estatales para desarrollar el levantamiento de datos, lo que incluye por lo general la gratificación de Aplicadores, Coordinadores Regionales y Supervisores que se contratan eventualmente para el efecto, así como una pequeña partida de insumos.
- ◆ El envío a cada una de las 32 entidades federativas de un Representante del INEE para apoyar al Coordinador Estatal y a su equipo de trabajo, en aquellas acciones que en lo general y en lo particular sean necesarias para alcanzar los mejores resultados del proceso.

5.4 Económicos. *Recursos disponibles para el operativo*

Según se mencionó en los incisos anteriores, las características técnicas, geográficas y las condiciones político-organizativas del Sistema Educativo Estatal, hacen necesario prever una serie de lineamientos y acciones apropiadas para realizar de mejor forma el levantamiento de datos.

Aunque a su vez, las condiciones generales de los levantamientos de datos, obedecen también a la disponibilidad de recursos económicos para ello, los cuales son casi siempre ajustados e imponen restricciones al momento de definir los modos de operar el proceso de aplicación. En otras palabras, un procedimiento o solución óptima debe ser económicamente factible

En ciertas ocasiones un procedimiento técnicamente deseable tiene que ser modificado por cuestión de costos. Por ejemplo, tal vez sea muy conveniente que el aplicador acuda a la escuela en la que trabajará a fin de que conozca con suficiente anticipación la vía de acceso, al mismo tiempo le brinde información al director y haga una inspección general del espacio en el que se realizará el llenado de instrumentos. Dado que esta sola acción duplica

los costos por concepto de aplicación (gratificación, traslados y pernoctas), las restricciones presupuestales han obligado a buscar soluciones alternas, con la consecuencia de que el encuentro del Aplicador con la autoridad escolar ocurre en un alto porcentaje hasta el día mismo de la aplicación, disminuyendo posibilidades de comunicación previa respecto al proceso de la aplicación, y que eso se haga mediante folletos.

A continuación se presentan algunas otras actividades del proceso que pueden ocurrir de manera diferente según los recursos con que se cuente:

- ◆ La capacitación directa a los aplicadores o a través de una estrategia de multiplicación “en cascada”
- ◆ La capacitación presencial o a distancia
- ◆ El tamaño del grupo de aplicación
- ◆ La atención del Aplicador de varias actividades simultáneas del proyecto al momento de su estancia en la escuela o contratación de varios aplicadores
- ◆ Ampliar la aplicación a varios días o intensificar las actividades en uno sólo.
- ◆ Cobertura del monitoreo de la calidad de la aplicación
- ◆ Monitoreo del proceso con personal interno o contratado
- ◆ La validación de la muestra de forma directa, a través de la estructura educativa o en forma automatizada empleando la información disponible.
- ◆ Envío y recolección de materiales de aplicación por “rutas de cordillera” o con un servicio de mensajero particular para cada entidad del País.
- ◆ Coordinación centralizada o por regiones en el estado.

Además de la disponibilidad de los recursos, debe considerarse el origen y aplicación de los mismos. En los levantamientos de datos de las evaluaciones que realiza el INEE, es éste quien aporta y ejerce los recursos necesarios para la reproducción, empaquetado, distribución y recolección del material. Asimismo, asigna a las AEE para que sean estas quienes manejen los recursos destinados a cubrir los gastos de traslado, viáticos y gratificación de aplicadores. Por otra parte, las AEE absorben ciertos costos indirectos por concepto de teléfono, uso de instalaciones y almacenes locales, capacitación de aplicadores, personal de apoyo, entre otros.

Es importante que el proceso de organización y conducción de la aplicación considere el registro de costos y la definición de indicadores de este tipo, a fin de que las modificaciones y mejoras del proceso se basen en un análisis de costo beneficio.

5.5 Éticos. *Orientaciones valorales de la organización*

Las anteriores consideraciones no tendrían un gran impacto externo si desde el interior de la organización que promueve las acciones, se maneja sin un sentido de ética y responsabilidad sobre las propias, por lo que en los levantamientos de datos, las normas y con-

sideraciones de calidad que persiguen traducirse en estándares se apegan a una serie de orientaciones valorales propias del Instituto (INEE, 2006, p. 2-6) y que al referirlas al proceso en cuestión se centran en las siguientes:

- ◆ El personal desarrollará sus actividades conforme con lo establecido en los manuales técnicos y operativos, así como en el modelo de calidad del INEE y en los lineamientos de organismos y asociaciones profesionales pertinentes.
- ◆ Se cuidará, con todo esmero, el manejo de datos y el control de los estándares de calidad en la aplicación de evaluaciones, y se evitará en todo momento que se descuiden o relajen por acción u omisión alguna.
- ◆ Llevar a cabo una política explícita sobre las evaluaciones y su uso, que vayan acordes, tanto con los lineamientos del INEE, como los establecidos internacionalmente.
- ◆ Actuar con el debido respeto a las personas e instituciones involucradas en los procesos de evaluación y esforzarse por evitarles cualquier daño o perjuicio.
- ◆ Proporcionar a las personas e instituciones que participan en los procesos de evaluación una información clara y adecuada sobre la naturaleza de la misma, así como de los derechos y responsabilidades de las personas que son evaluadas, y los principios éticos y las disposiciones legales que regulan las evaluaciones.
- ◆ Desarrollar y aplicar instrumentos y procedimientos que garanticen un manejo correcto de los datos, así como una interpretación correcta, precisa y clara de los resultados
- ◆ Vigilar la seguridad de los materiales de evaluación reservados o confidenciales, tanto en su almacenaje, como en el control de acceso a éstos, para garantizar que no queden expuestos públicamente de modo que su utilidad sea deteriorada o invalidada.
- ◆ Evitar todo tipo de acción u omisión que altere los resultados en cualquier sentido
- ◆ Custodiar y cuidar la documentación e información que conserve bajo su cuidado o a la cual tenga acceso, impidiendo o evitando su uso, sustracción, destrucción, ocultamiento o inutilización indebidos.
- ◆ El área de relaciones nacionales, difundirá información sobre las evaluaciones, antes y después de realizarlas para promover el interés y la cooperación de las personas involucradas.

6. EVIDENCIAS DE LA CALIDAD DE LOS LEVANTAMIENTOS DE DATOS COORDINADOS POR EL INEE

Como se señaló anteriormente y según se describe en la tabla 4, el INEE ha organizado y conducido 15 procesos de levantamiento de datos, que incluye piloteos y aplicaciones definitivas, de evaluaciones nacionales y otras internacionales.

Tabla 4. Aplicaciones coordinadas por el INEE, 2003-2006

Prog.	Fecha	Aplicación	Tipo de estudio		Población objetivo*			
			Aplicación nacional	Piloteo	Docentes	Alumnos	Escuelas	Nivel educativo
1	May-06	Segundo estudio regional comparativo y explicativo del laboratorio latinoamericano SERCE-LLECE	1		494	11,361	176	3 y 6° primaria
2	May-06	Condiciones y Prácticas Docentes en Primaria**	1		6,141	10,540	1,054	primaria (todos los grados)
3	May-06	Exámen de la calidad y el logro educativos Excale	1		4,469	64,518	3,482	3° primaria
4	Mar-06	Programa internacional para la evaluación del estudiante PISA	1		1,159	40,737	1,159	3° secundaria y 1° de bachillerato
5	Ene-06	Condiciones y Prácticas Docentes en Primaria**		1	1,899	3,940	394	1°, 3° y 5° primaria
6	Ene-06	Exámen de la calidad y el logro educativos Excale		1	2,483	8,600	354	3° primaria
7	Sep-05	Segundo estudio regional comparativo y explicativo del laboratorio latinoamericano SERCE-LLECE		1	176	3,489	98	4° primaria 1° secundaria
8	May-05	Exámen de la calidad y el logro educativos Excale	1		39,701	125,231	5,254	6° primaria 3° secundaria
9	May-05	Comparativa de estándares nacionales	1		0	10,002	448	6° primaria 3° secundaria
10	Abr-05	Programa internacional para la evaluación del estudiante PISA		1	100	1,800	54	3° secundaria y 1° de bachillerato
11	Feb-05	Pruebas Nacionales		1	371	8,439	318	6° primaria 3° secundaria
12	Jun-04	Programa de escuelas de calidad	1		10,062	135,926	946	2°, 3°, 4° y 5° primaria
13	Abr-04	Pruebas Nacionales	1		7,570	91,000	4,489	6° primaria 3° secundaria
14	Mar-03	Programa internacional para la evaluación del estudiante PISA	1		ND	35,730	1,157	Jóvenes de 15 años
15	Jun-03	Pruebas Nacionales de aprovechamiento en lectura y matemáticas	1		ND	266,046	4,716	3°, 5° y 6° de primaria y 1°, 2° y 3° de secundaria
Totales			10	5	74,655	817,359	24,099	

Las evaluaciones nacionales que se han realizado a partir de 2005 están basadas en un plan general de largo plazo que considera la medición escalonada en grados clave, esto es tercero de preescolar, tercero y sexto de primaria, tercero de secundaria y tercero de educación media superior, según se ilustra en la tabla que aparece a continuación. Cada grado se evalúa con intervalos de cuatro años, además en sexto de primaria y tercero de secundaria se hacen aplicaciones intermedias en muestras pequeñas, con representatividad nacional, para tener mayor precisión en las tendencias.

Tabla 5. Plan General de Evaluación

	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016
00			H	IP			H	IP			H	IP
03		ENMS				ENMS				ENMS		
06	EM	IP	EM	IP	ENMS		EM	IP	ENMS		EM	IP
09	EM			ENMS		EM		ENMS		EM		ENMS
12				EM	IP		ENMS	IP				ENMS
I		PISA SERCE			PISA		SERCE	PISA			PISA	SERCE

00= 3° Preescolar; 03= 3° Primaria; 06= 6° Primaria; 09= 3° Secundaria; 12= 3° Media Superior; E= Español; M= Matemáticas; N= Ciencias Naturales; S= Ciencias Sociales; H= Habilidades; I= Internacionales; E= Especiales; IP= Insumos y Procesos

Este diseño ha permitido reducir considerablemente el número de grados, escuelas y alumnos a evaluar, sin detrimento del propósito de conocer los aprendizajes logrados por los alumnos en el sistema educativo, antes bien éste se ha fortalecido ya que posibilitará comparaciones longitudinales y por la conformación de las muestras se tendrá representatividad nacional y por entidad, así como de los principales tipos de servicio y materias.

El diseño de largo plazo considera también la complementariedad de los estudios internacionales con los nacionales, así como la articulación entre estudios de resultados con otros de insumos y procesos que se han iniciado a partir de 2006.

Adicionalmente, se han establecido como requisitos que todos los instrumentos a utilizar sean previamente piloteados y que tanto las aplicaciones definitivas como las de los pilotos se hagan en fechas similares, para entre otras cosas sean actividades previstas en el sistema educativo, no obstante que son pocas las escuelas participantes.

Por otra parte, un dato objetivo que da cuenta de la calidad del levantamiento de datos es la proporción de instrumentos aplicados respecto a los considerados en la muestra, el cuadro siguiente muestra un análisis comparativo de los niveles logrados en diferentes estudios.

Tabla 6. Porcentajes de aplicación alcanzados en diversos estudios

Aplicación	Escuelas			Alumnos		
	Muestra	Aplicadas	%	Muestra	Aplicadas	%
PISA 2003	1 157	1 124	97.1	35 730	29 983	83.9
PISA 2006	1 159	1 150	99.2	40 737	33 766	83.0
PN 2004	3 294	3 166	96.0	59 131	53 581	91.0
Excale 2005	2 855	2 779	97.0	58 953	52 586	89.0
Excale 2006	3 357	3 189	95.0	62 600	55 436	89.0
SERCE	992	922	92.9	9 918	7 915	80.0
CPDP	175	60	91.0	11 361	9 584	84.4

Como se puede ver, aunque en todos los casos el porcentaje de escuelas aplicadas está por arriba del 90, las proporciones se reducen al considerar la cobertura de alumnos alcanzada.

Los porcentajes de aplicación varían de manera considerable al hacer la estimación por entidad o por estrato del estudio, como lo muestran los dos cuadros siguientes, respectivamente; hay casos críticos en los que las bajas proporciones de datos recolectados respecto a lo esperado según la muestra pone en riesgo la representatividad de la unidad de análisis.

Tabla 7. Porcentaje de aplicación en alumnos logrado en Excale 2006 por entidad.

Entidad	Muestra inicial	Aplicadas	Porcentaje de aplicación
Aguascalientes	1947	1871	96%
Baja California	2237	2049	92%
Baja California Sur	1977	1815	92%
Campeche	1754	1584	90%
Coahuila	1964	1807	92%
Colima	1850	1728	93%
Chiapas	2603	1655	64%
Chihuahua	2493	2035	82%
Distrito Federal	1559	1496	96%
Durango	2206	1778	81%
Guanajuato	1667	1526	92%
Guerrero	1936	1603	83%
Hidalgo	2108	1880	89%
Jalisco	1902	1737	91%
México	1891	1736	92%
Michoacán	2461	1904	77%
Morelos	2270	2105	93%
Nayarit	2006	1611	80%
Nuevo León	1884	1692	90%
Oaxaca			
Puebla	2389	2197	92%
Querétaro	1835	1744	95%
Quintana Roo	2118	2004	95%
San Luis Potosí	1963	1795	91%
Sinaloa	1900	1685	89%
Sonora	1922	1795	93%
Tabasco	1598	1459	91%
Tamaulipas	1940	1797	93%
Tlaxcala	2137	1988	93%
Veracruz	2115	1821	86%
Yucatán	2366	2157	91%
Zacatecas	1602	1382	86%
Totales	62600	55436	89%

Tabla 8. Porcentajes de aplicación alcanzados en dos estudios según estratos de escuelas

Estudio	Estrato	Alumnos		
		Muestra	Aplicadas	%
Excale 2006	Urbana pública	24 988	23 607	94.5
	Urbana privada	13 694	12 850	93.8
	Rural pública	14 978	12 823	85.6
	Indígena	7 064	5 407	76.6
	Cursos comunitarios	1 876	747	39.8
CPDP 2006	Completa	5 478	5 365	98.0
	Multigrado	2 430	1 995	82.0
	Cursos comunitarios	2 010	555	28.0

Es importante aclarar que esta pérdida de datos es imputable al trabajo de campo, pero a esto debe sumarse lo que se pierde en la etapa de lectura o captura, según sea el caso.

El hecho de que la aplicación no pueda ocurrir en escuelas completas se debe generalmente a que se trata de una escuela diferente a la considerada en la muestra o a que no hubo actividades escolares ese día, por múltiples motivos.

Estos incidentes pueden evitarse si la muestra de escuelas seleccionadas se valida previamente para corregir su información, si fuera el caso, o bien proceder a reemplazarla por otra, en caso de ser necesario.

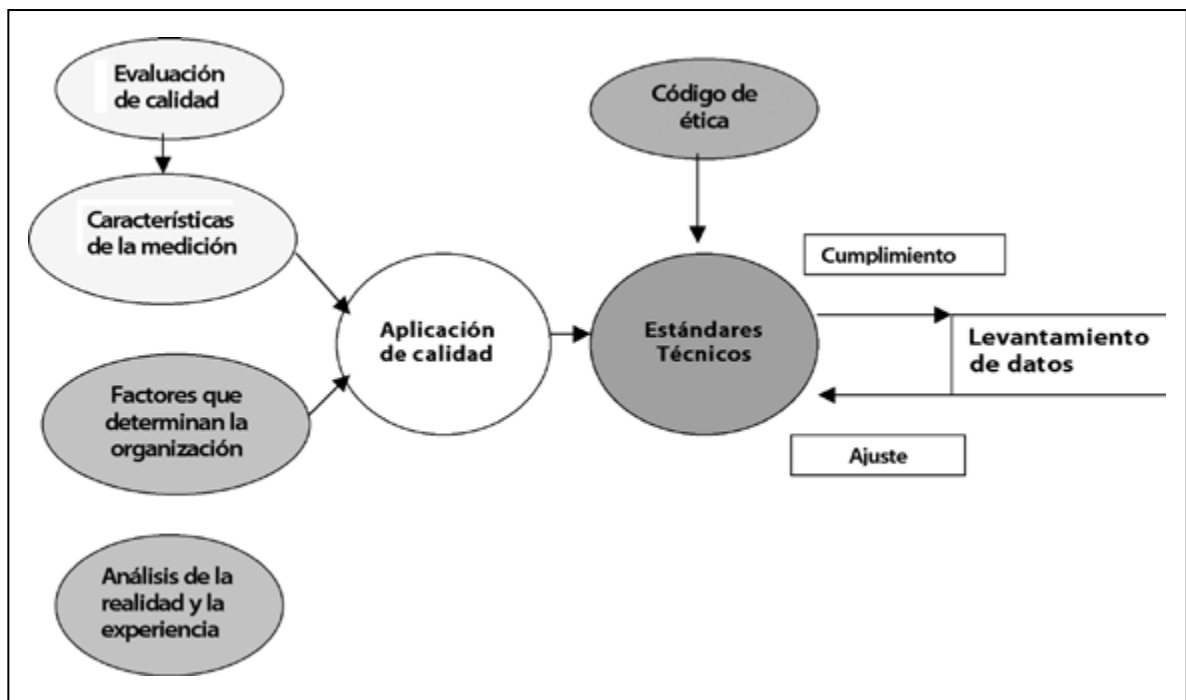
Otra medida que complementa la anterior es que se debe notificar al plantel participante, invariablemente y con suficiente anticipación, para evitar que programen otras actividades el día de la aplicación.

7. METODOLOGÍA SEGUIDA EN LA DEFINICIÓN DE LOS ESTÁNDARES TÉCNICOS PARA LOS LEVANTAMIENTOS DE DATOS

La propuesta de construir estándares para apoyar la compleja actividad del levantamiento de datos, ha tenido un camino de reflexión sobre las etapas del proceso general y de alguna manera transversal sobre los procedimientos del mismo.

Una idea general sobre su desarrollo se presenta en la siguiente figura:

Figura 1. Esquema del proceso seguido para la definición de estándares técnicos



Un análisis de la realidad y la experiencia llevada a cabo en las diversas aplicaciones antes señaladas, en contraste con los indicadores de calidad de los estudios de evaluación del INEE, propiciaron generar un modelo de calidad para el proceso de levantamiento de datos.

En otro momento, las partes de ese modelo, denominadas *consideraciones de calidad para las aplicaciones* han tratado de enfocarse y analizarse, en contraste con las etapas del proceso general, distinguiendo aquellas características puntuales de las mismas que son de vital importancia para asegurar dicho modelo de calidad.

La descripción de los estándares ha tenido dos configuraciones al considerar su nivel de riesgo sobre su incumplimiento, catalogando a los de más alto riesgo como estándares de calidad y a los subsecuentes como estándares de soporte.

Así mismo, la definición de los mismos ha tenido una serie de principios orientadores que buscan dar la mayor efectividad de su uso, entre ellos:

- ◆ una redacción clara y con matices de objetivación
- ◆ un balance o mediación de los alcances de los mismos al no ser de un carácter en extremo general o en extremo específico
- ◆ la idea de ser perfectibles pero con la firme intención de su cumplimiento, lo que denota el uso de los verbos tener y deber y finalmente
- ◆ con el sentido de una acción sinérgica, es decir, con interconexión entre ellos y por otro lado abarcando todos los puntos nodales de cada procedimiento.

A fin de asegurar la mayor objetivación de los estándares, la descripción general deriva o se apoya en una serie de descriptores que hacen referencia a las posibilidades del cómo tener evidencia de los mismos y a una precisión o aproximación a ciertos marcadores de tiempo o cantidad para su cumplimiento.

Finalmente la metodología de construcción prevé la imperiosa necesidad de tener un seguimiento sobre la difusión, uso y retroalimentación de los mismos, con espacios específicos de consulta, pero también con ejercicios de monitoreo sobre su aplicación.

8. NOCIÓN Y CARACTERÍSTICAS DE UNA APLICACIÓN DE CALIDAD

El tema de la calidad supone en nuestros días, más que una idea simple de adjetivo, un concepto elaborado que encierra en si mismo una serie de categorías bien definidas y conectadas entre sí.

La evaluación educativa, que es la *emisión de un juicio como resultado de comparar una realidad empírica (...) con un parámetro normativo previamente definido ...* (INEE 2003, p.3) lleva en si mismo una idea de calidad que constantemente es referida. Una evaluación con calidad debe ser:

- ◆ Confiable. Pues cuenta con validez y solidez técnica e incluye enfoques diversos y complementarios
- ◆ Define parámetros pertinentes para comparar los resultados de las mediciones
- ◆ Ofrece juicios medidos a partir de la comparación entre mediciones y parámetros
- ◆ Tiene difusión amplia, oportuna y transparente (INEE, 2003)

Los procesos de las grandes evaluaciones o evaluaciones estandarizadas tienen sin embargo varios momentos o fases de ejecución, cada una de las cuales tiene su propio proceso de calidad.

El *levantamiento de datos* y en específico *la aplicación*, son actividades de desarrollo en campo que trascienden al realizarse en forma controlada y asegurar de ese modo la igualdad de condiciones de los sustentantes, (INEE, 2005) contribuyendo de esta manera a la confiabilidad de los resultados.

El INEE considera que una aplicación de calidad debe apegarse a las siguientes consideraciones:

- ◆ se realiza con procedimientos estandarizados en todas las escuelas participantes
- ◆ con respeto a las personas evaluadas y a la organización escolar
- ◆ con información suficiente y oportuna a los involucrados
- ◆ con medidas de seguridad para resguardar la integridad y confidencialidad del material
- ◆ con una mínima proporción de instrumentos no contestados
- ◆ con bajos índices de irregularidades durante la aplicación y
- ◆ con eficiencia en el manejo de recursos.

Los esfuerzos actuales en materia de levantamiento de datos están encaminados a observar los anteriores indicadores.

Al respecto, el área responsable dentro del Instituto² de coordinar el proceso de aplicación expresa su misión en la siguiente forma:

Instrumentar los procesos de levantamiento de datos con estrategias innovadoras y de alta calidad técnica que reduzcan la posibilidad de sesgos atribuibles a las condiciones del proceso mismo.

Vislumbrando la esencia del objetivo y recordando los factores de la organización que se han descrito en el anterior punto 5, las condiciones en las que ha de lograrse el mismo previenen un gran reto: Conducir una aplicación estandarizada en un entorno heterogéneo.

9. ESTÁNDARES TÉCNICOS PARA EL LEVANTAMIENTO DE DATOS

La exposición de los estándares técnicos conlleva enunciar las normas principales a las que se sujeta el proceso de levantamiento de datos, buscando con ello contribuir a tener una actividad ordenada, clara y con la cooperación de todos los involucrados.

Por el nivel de desagregación dentro de todo el proceso general, los estándares técnicos pueden considerarse estándares de especificación, pues persiguen evitar errores de identificación en las tareas a realizar, con un lenguaje claro y preciso.

Los estándares técnicos que a continuación se expresan están en concordancia con los criterios de calidad que se han definido para el proceso de levantamiento de datos, mismos que a su vez pueden considerarse indicadores de medición. De igual forma, dichos estándares se han definido con la perspectiva de estar detallados bajo el esquema de descriptores, formulas o indicadores cualitativos, así como su identificación exacta con evidencias documentales en las distintas etapas del proceso.

Como se dijo anteriormente, se distinguen dos grandes tipos de estándares: 1) estándares de calidad, que se aprecian como de aplicación obligatoria y 2) estándares de soporte,

² Subdirección de Logística en Evaluación Educativa, dependiente de la Dirección de Relaciones Nacionales.

que tienen más un carácter de recomendación. En este texto se distinguirán los primeros por resaltarse en negritas.

A continuación se enuncia la parte medular de los mismos.

Aplicación normalizada

1. Las condiciones en donde se desarrolla la aplicación de los instrumentos de una evaluación, deben ser lo mas homogéneas posibles, en sus tiempos, formas y espacios
 - 1.1. Los procedimientos de aplicación deberán ser uniformes en tiempo, forma y espacio en la totalidad de las escuelas
 - 1.2. Todos los procedimientos de aplicación deben contar con manuales técnicos de operación claramente redactados, con un carácter práctico y de prevención sobre las actividades a desarrollar y sin contradicciones entre sus partes o con otros similares
 - 1.3. En la contratación de Aplicadores se deberá considerar uno por cada grupo de aplicación, a fin evitar desigualdades o deficiencias en la administración de instrumentos
2. Las personas participantes directamente en el proceso de aplicación deberán estar debidamente capacitadas en su totalidad
 - 2.1. Todos los Aplicadores seleccionados deben siempre recibir la capacitación correspondiente antes de efectuar su función
 - 2.2. La capacitación deberá desarrollarse con base a las líneas generales propuestas por el INEE a través de sus materiales de apoyo operativo

Respeto a las personas y a la organización escolar

3. Los evaluadores deben conducirse en todo momento con una actitud de respeto con quienes interactúen en el proceso
 - 3.1. Los alumnos y demás personal participante en el estudio de evaluación deberá recibir un trato formal y con respuestas claras a todas sus inquietudes
 - 3.2. Será indispensable el uso del protocolo de presentación y despedida en cada centro escolar por parte de quienes acudan a una visita con motivo del estudio de evaluación
 - 3.3. Los Coordinadores Regionales, Aplicadores y Monitores de calidad participantes contarán invariablemente con un seguro contra accidentes durante el tiempo de su comisión, para lo cual el AEE deberá enviar con oportunidad los datos correspondientes al INEE
 - 3.4. La relación de trabajo y asesoría eventual que se de entre Coordinadores, Representante del INEE y Aplicadores deberá mantener siempre un trato formal

Información suficiente y oportuna

4. Es necesario que todos aquellos que se involucren en la aplicación, cuenten con la mayor cantidad de información posible sobre los propósitos y alcances del proyecto a fin de obtener una óptima colaboración de los participantes
 - 4.1. Los acuerdos de apoyo para los levantamientos de datos entre las Entidades y el INEE deberán realizarse en todo caso de manera formal mediante las invitaciones, convocatorias y respuestas respectivas
 - 4.2. En aquellos casos donde el INEE envíe un Representante como apoyo a la entidad, este deberá contar con el perfil y la información necesaria para desarrollar cabalmente su función definida en su Manual respectivo
 - 4.3. Las áreas estatales de evaluación deberán recibir 30 días previos a la aplicación, notificación formal de la muestra de escuelas que participarán en el estudio a fin de prever su logística interna de atención
 - 4.4. Las escuelas participantes deberán ser notificadas y recibir un folleto informativo 20 días previos a la aplicación
 - 4.5. Las áreas estatales de evaluación deberán recibir un cronograma general de trabajo donde se especifique entre otras fechas, la de entrega y recolección de material a fin de asegurar su correcta recepción y su posterior devolución al INEE
5. Los materiales de evaluación deben entregarse a las áreas estatales de evaluación en orden, cantidad y tiempo suficientes para organizar localmente su aplicación.
 - 5.1. Las entidades deberán recibir 20 días previos a la aplicación, el material de capacitación correspondiente
 - 5.2. Las entidades deberán recibir 10 días previos a la aplicación, el material de capacitación correspondiente

Seguridad del material

6. Debe asegurarse que los materiales de evaluación estén bajo un resguardo organizado desde su salida de la imprenta hasta su regreso a los laboratorios de lectura
 - 6.1. El embalaje donde se coloque el material de aplicación para cada grupo desde la imprenta, deberá sellarse con etiquetas que aseguren su inviolabilidad hasta su apertura en el centro escolar por parte del aplicador
 - 6.2. Las entidades participantes recibirán un paquete del material de evaluación por cada grupo de aplicación, mismo que deberá ir identificado con los datos de destino y cantidades de su contenido en etiquetas externas claras y visibles
 - 6.3. El material enviado del INEE al AEE será permanentemente monitoreado durante su traslado y se verificará la totalidad de su entrega

- 6.4. Los materiales de aplicación serán entregados a los aplicadores en el momento mas cercano a la fecha de aplicación, mismo que será definido según el tiempo que requiera para trasladarse a la escuela asignada
- 6.5. Los materiales de evaluación entregados al Aplicador deberán verificarse invariablemente, mediante el conteo individual de los mismos y se hará su registro en el formato correspondiente que haga llegar la Subdirección de Logística del INEE al área estatal de evaluación
- 6.6. El Aplicador asignado a cada escuela, es la persona indicada para trasladar los materiales de evaluación a la misma
- 6.7. Será responsabilidad del Aplicador el que los materiales de evaluación que administre en la escuela, no sean fotocopiados, copiados, mutilados o extraviados al momento de la aplicación
- 6.8. Es indispensable el uso de los formatos de control correspondientes por parte del Aplicador para asegurar el resguardo y recuperación de la totalidad de los materiales posterior a la aplicación
- 6.9. Los materiales aplicados deberán devolverse por el Aplicador al área estatal de evaluación en un tiempo máximo de 24 horas después de la evaluación
- 6.10. Los materiales aplicados deberán ser cuidadosamente revisados en cantidad y forma, así como puntualmente contrastados con los datos de entrega de los mismos, por medio de una estrategia específica definida por el área estatal de evaluación
- 6.11. La empresa contratada para la recolección del material aplicado en cada entidad, deberá contrastar el mismo con una guía de entrega y recepción previamente asignada a fin de garantizar la recuperación total

Proporción suficiente de escuelas e instrumentos aplicados

7. Es necesario que se valide de manera previa la información de las escuelas a participar, para mantener o aumentar los porcentajes de aplicación

- 7.1. La información de la muestra de escuelas a participar en el estudio deberá validarse previamente por las áreas estatales de evaluación, según los lineamientos y programas informáticos que le proporcione el INEE, con el objeto de disminuir al máximo la no aplicación
- 7.2. Deberá asegurarse que el porcentaje de escuelas aplicadas sea mayor al 85%
- 7.3. Deberá asegurarse que el porcentaje de participación de actores previstos para la evaluación sea mayor al 80%
- 7.4. De no alcanzarse las cifras anteriores en una entidad, se analizará cada caso para valorar la conveniencia de excluirse del informe de resultados. Si decide incluirse, deberá hacerse la indicación correspondiente respecto a la no representatividad

Porcentajes mínimos de Irregularidades

8. Es muy recomendable que las personas que participan como Aplicadores, reúnan un perfil consistente en lo académico, social y profesional
 - 8.1. Los Aplicadores contratados deberán tener un perfil que solvete habilidades comunicativas en ambientes escolares, de respeto y desarrollo profesional eficiente con el fin de evitar irregularidades durante la aplicación
 - 8.2. Los Aplicadores deberán contar con los oficios, cartas y croquis de rutas de acceso correspondientes para apoyarles en su actividad formal en la escuela
 - 8.3. La capacitación impartida a los Aplicadores deberá desarrollarse y verificarse con criterios de eficiencia, asegurándose de la total comprensión de conceptos y procedimientos clave del momento de aplicación
9. Se requiere de verificar con antelación que los puntos nodales de todos los procedimientos funcionen debidamente a fin de evitar irregularidades imputables al proceso mismo
 - 9.1. El material a aplicar que sea susceptible de lectura óptica, deberá monitorearse al momento de su proceso de impresión a fin de asegurar su posterior funcionamiento correcto. Deberá hacerse al menos una prueba por cada forma óptica
 - 9.2. El material de evaluación deberá tener durante su etapa de impresión un proceso de monitoreo del 2% para evitar errores de montaje o encuadernación
 - 9.3. Los responsables del empaquetado del material deberán contar con instrucciones claras y suficientes para dicha actividad, además de un monitoreo presencial para evitar errores de ordenamiento.
 - 9.4. El empaquetado de los materiales por grupo de aplicación deberá contar con un cuidadoso monitoreo, con una muestra aleatoria del 2%, para asegurar su correcta organización en la caja correspondiente
 - 9.5. Los procesos generales de aplicación, sean piloteos o definitivos, contarán con un monitoreo de la calidad en campo coordinado por el INEE, correspondiente al 2 % de escuelas seleccionadas al azar, a fin de obtener evidencias sobre los procedimientos estandarizados de los mismos y su respectiva contribución a la confiabilidad
 - 9.6. En cada grupo de aplicación se hará un reporte donde quede evidencia de las incidencias e irregularidades ocurridas durante el tiempo de aplicación. El INEE definirá el formato de dicho reporte
 - 9.7. Las actas de irregularidades levantadas en las escuelas, serán la base para integrar un informe que dé cuenta de las desviaciones del proceso, a nivel nacional. Esto deberá incluirse en el reporte técnico

Eficiencia de los procesos

10. Es necesario instrumentar medidas de seguimiento sobre el costo del impacto y la eficiencia con que se desarrolla cada proceso general de levantamiento de datos.
- 10.1. A fin de asegurar los recursos económicos a utilizar por las áreas estatales de evaluación en la etapa de aplicación, es necesario que estas realicen un presupuesto de costos basados en una estrategia de operación a nivel local y apegados a los lineamientos emitidos por el INEE
- 10.2. El INEE deberá notificar de manera formal a las áreas estatales de evaluación sobre el monto presupuestal autorizado para desarrollar la aplicación
- 10.3. Se harán registros puntuales de corte transversal y análisis longitudinales sobre los costos de impresión, traslado de material, envío de Representantes, aplicación y los monitoreos de calidad respectivos, para valorar de manera objetiva la eficiencia general de cada proceso

10. CONSIDERACIONES FINALES: SOBRE EL USO DE LOS ESTÁNDARES

Finalmente, a manera de conclusión se presentan a continuación algunas consideraciones generales que tienen que ver particularmente con la consulta y uso de los estándares.

1. Estos estándares fueron definidos a partir de la experiencia de las aplicaciones que ha realizado el INEE y para ser utilizados en futuros procesos de levantamientos de datos de sus evaluaciones, aunque obviamente pueden servir de guía para otros estudios de características similares.
2. El esfuerzo de definir estándares resulta estéril si no se discuten, adoptan, y utilizan.
3. El personal del INEE que se encarga de organizar y coordinar los procesos de levantamiento de datos son los principales destinatarios de esta propuesta, aunque se extiende también al personal de apoyo del Instituto, al personal de las Áreas Estatales de Evaluación, directores y profesores de las escuelas evaluadas y todos aquellos que participan como aplicadores, monitores de calidad o coordinadores regionales.
4. Estos estándares no son todos, ni son estáticos, seguramente a partir de su uso, la retroalimentación que del trabajo de campo surja, y sobre todo con base en los reportes de irregularidades que se obtengan, se atenderá a una complementación y actualización permanente de estos y deberán ser sometidos a la aprobación periódica del Consejo Técnico del INEE.
5. Los estándares son orientaciones para hacer mejor el levantamiento de datos y valorar el proceso, es claro que difícilmente se podrá cumplir con todos desde la primera vez, aunque no por ello se deberá calificar como malo el proceso. Los estándares deberán ser

una aspiración permanente de todos los participantes en el proceso.

6. Otra aplicación de los estándares técnicos es que constituyen un referente concreto para elaborar el reporte técnico del trabajo de campo, para valorar el proceso en general y en procedimientos específicos, a la vez que para establecer puntos de comparación y mejora entre un operativo y otro.

11. REFERENCIAS BIBLIOGRÁFICAS

- BROWN, FREDERICK G. (2000). Principios de la medición en psicología y educación. (Traducción del inglés). Manual Moderno. México.
- BACKHOFF E., L. MONROY, M. PEÓN, A. SÁNCHEZ Y M DE L. TANAMACHI. (2005). Exámenes de la calidad y el logro educativos (Excale): nueva generación de pruebas nacionales. INEE. Colección de cuadernos de investigación. México.
- GAJARDO MARCELA –DIRECTORA- (2002). Preal. año 3 num. 13. En internet: www.preal.org/acceso, agosto 2004
- INECSE (2004). Marcos teóricos de PISA 2003. Conocimientos y destrezas en Matemáticas, Lectura, Ciencias y Solución de problemas. INECSE. Ministerio de Educación y Ciencia. España. 226 pp.
- INEE (2003). La calidad de la educación básica en México. Folleto de difusión. P. 3. INEE/ Fundación Este país. México
- INEE (2005). ¿Qué es el INEE? Sus fines y estrategias al servicio de la sociedad. col. Temas de la evaluación num. 6. INEE. México.
- INEE. (2006). Código de conducta. INEE, México, 13 pp.
- JOINT COMMITTEE ON STANDARDS FOR EDUCATIONAL EVALUATION (JCSEE). 1988. Normas de evaluación para programas, proyectos y material educativo. Ed Trillas. Mex.
- MARTÍNEZ RIZO F. EDUARDO BACKHOFF E., OTROS. (2000). Estándares de calidad para instrumentos de evaluación educativa. Ceneval. México. 52 pp.
- MARTÍNEZ RIZO F. (2004). La comparabilidad de los resultados de las pruebas nacionales de lectura y matemáticas, 1998-2003. INEE. Colección de cuadernos de investigación. Mex
- MARTÍNEZ RIZO F. (2005). La evaluación educativa en México. Mimeo
- ORELAC, UNESCO, SEP. (2005) Panorama educativo 2005: progresando hacia las metas. Proyecto regional de indicadores educativos. Cumbre de las Américas. SEP-ORELAC. México, 111 pp.
- SEP (2000) Memoria del Quehacer educativo. Vol I. Secretaría de educación pública. México. 438 pp.
- SEP (2001). Programa Nacional de Educación. Secretaría de educación pública. México. 269 pp.

- SEP. (2002). La Experiencia de la Dirección General de Evaluación en la Educación Básica y Normal: 30 años de medición del logro educativo. México.
- SEP (2006). Estadísticas. En internet:www.sep.gob.mx Acceso, agosto 2006.
- UNESCO (2003). La experiencia del proyecto regional de indicadores educativos 2000 2003. Cumbre de las Américas. Proyecto regional de Indicadores Educativos. Santiago de Chile. UNESCO. Chile. 89

DIRECTORIO DE LA ÁREAS ESTATALES DE EVALUACIÓN

- AGUASCALIENTES.** Josefina Mercado Haro
- BAJA CALIFORNIA.** José Manuel González Osuna
- BAJA CALIFORNIA SUR.** Ignacio Alfredo Higuera
- CAMPECHE.** Miriam Cecilia Lira Valencia
- COAHUILA.** Guillermo Ortiz Vázquez
- COLIMA.** Andrés Magaña Moreno
- CHIAPAS.** Armando Rojas Hernández
- CHIHUAHUA.** Jesús Eliseo Ríos Durán
- DISTRITO FEDERAL.** Maria de Lourdes Atilano Vireles
- DURANGO.** Marcial Antúnez Bautista
- GUANAJUATO.** Rafael Díaz García
- GUERRERO.** Ildegunda Santos Jiménez
- HIDALGO.** Ernesto Ordaz Labra
- JALISCO.** Israel Jacobo Bojorquez
- MÉXICO.** Elizabeth Hernández González
- MICHOACÁN.** Salvador Cervantes Herrera
- MORELOS.** Olga Yolanda Valdéz Zermeño
- NAYARIT.** Mirtha Evelia Mares López
- NUEVO LEÓN.** Alfredo Fernández Domínguez
- OAXACA.** José Raymundo Martínez Morga
- PUEBLA.** José Arturo Ramírez y López
- QUERÉTARO.** Ma. Del Pilar Puga Tovar
- QUINTANA ROO.** Enrique Eduardo Encalada Sánchez
- SAN LUÍS POTOSÍ.** Rosa María García Martínez
- SINALOA.** Santiago Zúñiga Barrón
- SONORA.** José Luí Ibarra Apocada
- TABASCO.** Ma. Encarnación Guadalupe Tamés Cueto
- TAMAULIPAS.** Noemí González Barragán
- TLAXCALA.** Prisciliano Molina Atonal
- VERACRUZ.** Lourdes Fragoso López
- YUCATÁN.** Hortensia González Urzaiz
- ZACATECAS.** Arnulfo Pinedo Saldivar

TÍTULOS DE ESTA COLECCIÓN:

- 1.- *Los resultados de las pruebas de PISA*
Martínez Rizo, Felipe.
- 2.- *Factores externos e internos a las escuelas que influyen en el logro académico de los estudiantes de nivel primaria en México, 1998-2002*
Muñoz I, Carlos et al.
- 3.- *Contextualización sociocultural de las escuelas de la muestra de estándares nacionales (1998-2002).*
_____ *Determinantes sociales y organizacionales del aprendizaje en la Educación Primaria de México. Un análisis de tres niveles.*
_____ *Perfil de las escuelas primarias eficaces de México (2001)*
Fernández, Tabaré.
- 4.- *Tercer Estudio Internacional de Matemáticas y Ciencias Naturales (TIMSS): resultados de México en 1995 y 2000*
Backhoff, Eduardo y G. Solano.
- 5.- *Estudio Sobre las Desigualdades Educativas en México: la Incidencia de la Escuela en el Desempeño Académico de los Alumnos y el rol de los Docentes*
Treviño, Ernesto y G. Treviño.
- 6.- *Evaluación inicial de los procesos de calibración y equiparación de las pruebas del proyecto de estándares nacionales.*
Magriñá, Antonio.
- 7.- *Factores asociados al aprendizaje del lenguaje y las matemáticas en 13 estados de México.*
Cervini, Rubén.
- 8.- *Acciones de Evaluación en las Instituciones Públicas de Educación Media Superior*
Antonio, Rocío.
- 9.- *El proyecto PISA: su aplicación en México*
Vidal, Rafael, et. al.
- 10.- *La Comparabilidad de los Resultados de las Evaluaciones. Importancia y Dificultad de la Equiparación*
Martínez Rizo, Felipe.
- 11.- *Marginación y rezago educativo en México.*
Ávila, José Luis.
- 12.- *Pruebas y rendición de cuentas*
Martínez Rizo, Felipe.
- 13.- *Panorama Educativo 2004. La edición 2004 de Education at a Glance de la OCDE*
Martínez Rizo, Felipe.
- 14.- *El Diseño de Sistemas de Indicadores Educativos: Consideraciones Teórico- Metodológicas*
Martínez Rizo, Felipe.
- 15.- *La telesecundaria mexicana: desarrollo y problemática actual*
Martínez Rizo, Felipe.
- 16.- *Sobre la difusión de resultados por escuela*
Martínez Rizo, Felipe.
- 17.- *Exámenes de la Calidad y el Logro Educativos (Excale): nueva generación de pruebas nacionales*
Backhoff Escudero, Eduardo.
- 18.- *La Educación Mexicana en Education at a Glance 2005*
Martínez Rizo, Felipe.
- 19.- *Metodología para evaluar la calidad de las traducciones de las pruebas internacionales: el caso mexicano de TIMSS-1995*
Guillermo Solano-Flores, Eduardo Backhoff Escudero y Luis Ángel Contreras-Niño

- 20.- *Acerca de la Validez de los Exámenes de la Calidad y el Logro Educativos (Excale)*
María Araceli Ruiz-Primo, Jesús M. Jornet Meliá y Eduardo Backhoff Escudero
- 21.- *Pruebas de selección y pruebas para evaluar escuelas: nuevas consideraciones sobre su uso y la difusión de sus resultados*
Martínez Rizo, Felipe.
- 22.- *Promedio de escolaridad y nivel de desigualdad de años de escolaridad en las entidades y municipios de la República Mexicana*
Adán Moisés García Medina (con la colaboración de Edna Huerta Velásquez)
- 23.- *Las primarias comunitarias y su desempeño Consideraciones a partir del estudio comparativo 2000-2005 del INEE*
Felipe Martínez Rizo