

DISEÑO Y DESARROLLO DE LOS EXÁMENES DE LA CALIDAD Y EL LOGRO EDUCATIVOS

EDUARDO BACKHOFF / ANDRÉS SÁNCHEZ / MARGARITA PEÓN /

LUCÍA MONROY / MARÍA DE LOURDES TANAMACHI

Resumen:

Se describe la puesta en marcha de una nueva generación de pruebas nacionales que el Instituto Nacional para la Evaluación de la Educación (INEE) utiliza, en muestras de estudiantes con representatividad nacional y estatal, para evaluar la calidad del sistema educativo nacional con base en los aprendizajes logrados en educación básica y media superior. Se describen los propósitos y principios de la evaluación del aprendizaje que asume el INEE, así como el marco de referencia de los Exámenes de la Calidad y el Logro Educativos. Se definen las características de las pruebas nacionales como: criterios, alineadas al currículum nacional y con un diseño matricial. Se detalla el proceso para diseñar, construir, aplicar y validar estos exámenes. Se concluye discutiendo la necesidad de contar con instrumentos evaluativos de calidad para cumplir con las metas del INEE y así ayudar a mejorar la educación en México.

Abstract:

The article describes the initial use of a new generation of national tests by Mexico's institute of educational evaluation (Instituto Nacional para la Evaluación de la Educación—INEE) among sample groups of students who are representative in national and state terms. The aim is to evaluate the quality of the national education system, based on achievement attained in elementary and junior high school. A description is given of the objectives and principles assumed by INEE in evaluating learning, as well as the framework of reference of examinations of educational achievement and quality. The characteristics of national tests are defined: criteria, alignment with the national curriculum and the design matrix. The process is detailed for designing, constructing, applying and validating these examinations. The conclusion discusses the need to have evaluative instruments of quality to comply with the goals of INEE and thus help to improve education in Mexico.

Palabras clave: pruebas de aprovechamiento, calidad educativa, educación básica, México.

Key words: achievement test, educational quality, elementary education, Mexico.

Los autores están adscritos a la Dirección de Pruebas y Medición del Instituto Nacional para la Evaluación de la Educación (INEE): Andrés Sánchez es subdirector de Pruebas de Matemáticas y Ciencias Naturales; Margarita Peón es subdirectora de Pruebas de Español y Ciencias Sociales; Lucía Monroy y María de Lourdes Tanamachi son jefas de proyecto.

Eduardo Backhoff es director de Pruebas y Medición del INEE. José María Velasco núm. 101, col. San José Insurgentes, del. Benito Juárez, CP 03900, México, DF, CE: backhoff@inee.edu.mx

Hasta el año 2002 la Dirección General de Evaluación (DGE) de la Secretaría de Educación Pública (SEP) fue la única instancia gubernamental en educación básica responsable de evaluar al Sistema Educativo Nacional (SEN). Para ello, desde su creación, en la década de los años setenta, la DGE ha diseñado y operado un sinnúmero de programas e instrumentos evaluativos con el fin de conocer los resultados educativos de las escuelas primarias y secundarias del país (Velázquez, 2000; DGE, 2000).

Aunque no se hicieron públicos los resultados de estos programas ni todos tuvieron el impacto deseado, la DGE coadyuvó a poner en la mesa de las decisiones políticas el tema de la evaluación educativa como una vía indispensable para mejorar la calidad de la educación nacional; presencia que se vio reforzada, en gran medida, por el impacto de los resultados de aprendizaje de las evaluaciones internacionales, como es el caso de PISA (por sus siglas en inglés, *Programme for International Student Assessment*). Como una consecuencia de lo anterior, el Gobierno Federal en su Programa Nacional de Educación 2001-2006 (PRONAE) consideró la creación de mecanismos rigurosos y confiables de evaluación que operen con independencia de las autoridades y permitan rendir cuentas a la sociedad (SEP, 2001).

El PRONAE estableció dos grandes metas: *a*) la creación del Instituto Nacional para la Evaluación de la Educación (INEE) y *b*) la redefinición de las funciones de la DGE. El primero se concibió como un organismo con independencia y autonomía, cuyo objetivo primordial es conocer y explicar la calidad del SEN en su conjunto, así como dar a conocer los resultados de sus evaluaciones a las autoridades educativas y a la sociedad en general. Por su parte, la DGE se redefinió como una instancia de la administración central, cuyo propósito es realizar evaluaciones diagnósticas y formativas orientadas a retroalimentar la toma de decisiones sobre individuos e instituciones (DGE, 2004).

Con la idea de evaluar la calidad de la educación en México, la DGE desarrolló, desde hace más de un lustro, lo que ahora se conoce como las Pruebas de Estándares Nacionales (PEN). Dichos instrumentos fueron utilizados en el periodo 1998-2003, con muestras representativas de escolares de todo el país. Sin embargo, al crearse el INEE, estas pruebas pasaron a ser parte de su responsabilidad; función que asumió de manera parcial con el análisis de los resultados de la aplicación realizada –todavía por la DGE– en junio de 2003. Dichos resultados fueron publi-

cados en su primer informe *La calidad de la educación básica en México* (INEE, 2003).

Siguiendo el mismo esquema metodológico propuesto por la DGE en cuanto a estructura y contenidos, el INEE elaboró una nueva versión de las PEN, que aplicó en junio de 2004 con una muestra representativa de alumnos de todo el país, cuyos resultados se publicaron en su segundo informe anual *La calidad de la educación básica en México: resultados de evaluación educativa 2004* (INEE, 2004a).

Las PEN se diseñaron con la idea de evaluar las habilidades cognitivas más que los conocimientos. Para ello se desarrollaron varias pruebas de comprensión lectora y matemáticas, dirigidas a estudiantes de primaria y de secundaria. En el primer caso las evaluaciones se construyeron por grado escolar (2º, 3º, 4º, 5º y 6º); para secundaria se realizó una sola prueba dirigida a estudiantes de los tres grados.

Desafortunadamente, el diseño y construcción de las PEN adolecen de deficiencias técnicas importantes que hacen que no sea recomendable realizar comparaciones de los resultados en el tiempo; razón por la cual no es posible conocer las tendencias en los aprendizajes de los estudiantes. En publicaciones anteriores hemos señalado (Backhoff y Martínez-Rizo, 2004 y Backhoff, 2005) que entre estas deficiencias se encuentran las siguientes: *a)* las especificaciones de los contenidos no se encuentran bien definidas; *b)* no se diseñaron con una metodología *ex profeso* para su equiparación; *c)* los reactivos ancla no tuvieron la misma ubicación en las distintas versiones de las pruebas y *d)* ninguna evaluación se piloteó antes de su aplicación.

Por lo anterior, es imposible comparar en forma confiable los resultados de dos aplicaciones, como lo han demandado varios sectores de la sociedad (Martínez-Rizo, 2004). A estas deficiencias se suma el hecho de que la documentación en donde se plasma el marco de referencia que sustenta las pruebas es escasa, por lo que no queda claro si se realizaron los procesos mínimos que validan estos instrumentos.

Adicionalmente, las PEN no se diseñaron estrictamente con el propósito de evaluar los contenidos de los planes y programas de estudios nacionales, razón por la cual pueden decir poco acerca del dominio de los contenidos curriculares que poseen los escolares y, con esto, se limita la información que se le puede dar a las autoridades educativas sobre el logro de los propósitos del currículo mexicano (Backhoff y Martínez-Rizo, 2004; Backhoff, 2005).

Estas limitaciones son importantes si se toma en cuenta que entre los propósitos del INEE, que se relacionan con la evaluación del aprendizaje, se encuentran los siguientes: *a)* construir una visión general de lo que los estudiantes aprenden como resultado de su escolarización formal, *b)* conocer los puntos fuertes y débiles del aprendizaje de los estudiantes en las asignaturas de mayor importancia y *c)* permitir las comparaciones del rendimiento escolar, así como las tendencias de aprendizaje a lo largo del tiempo.

Tomando en cuenta las limitaciones de las PEN para lograr estos propósitos, el Consejo Técnico del INEE optó por la elaboración de un Plan General de Evaluación del Aprendizaje en el cual se considera la construcción de una nueva generación de pruebas nacionales, los Exámenes de la Calidad y el Logro Educativos (EXCALE), cuya preparación se inició en febrero de 2004 y su primera aplicación a nivel nacional se programó para junio de 2005 (INEE, 2004b).

A partir de lo anterior, los objetivos de este trabajo son: *a)* exponer los propósitos y principios del INEE en relación con la evaluación del aprendizaje; *b)* describir el marco de referencia de los EXCALE; *c)* describir en forma resumida el proceso de diseño, construcción, aplicación y validación de estas pruebas; y *d)* discutir la importancia de contar con instrumentos robustos para evaluar la calidad de los servicios educativos que se ofrecen en el país.

Los procedimientos reseñados en este artículo requerirían mucho más espacio para ser explicados y justificados detalladamente. Por ello, tras esta exposición general, se pretende hacer una serie de artículos posteriores que den cuenta detallada de los principales aspectos conceptuales y metodológicos.

Propósitos y principios de la evaluación del aprendizaje

En México la cobertura educativa del nivel básico se ha ido resolviendo gradualmente y, desde hace una década, el interés se ha centrado en indagar qué es lo que efectivamente aprende la población escolar y cuáles son los aprendizajes básicos que forman parte del repertorio de los estudiantes.

La evaluación del aprendizaje que realiza el Instituto Nacional para la Evaluación de la Educación tiene la intención de proporcionar al país un conocimiento general del rendimiento escolar que logran los estudiantes en las asignaturas básicas del currículum nacional; mismas que se definen como las propiamente instrumentales –español y matemáticas– y las

que cubren grandes áreas curriculares –ciencias naturales y sociales–, identificando en cada grado evaluado¹ la solución más idónea de acuerdo con el currículo del país (INEE, 2004b).

Sin embargo, la evaluación del aprendizaje puede ser entendida de muchas maneras, dependiendo de los propósitos para la que se utilice, los alcances que persigue, los objetos de la evaluación, los métodos que se empleen para realizarla así como las posturas teórico metodológicas en las cuales se fundamente. Con la idea de aclarar al lector los puntos de partida del INEE respecto de la evaluación del aprendizaje, a continuación se definen los propósitos que persigue dicha valoración y los principios de los que parte.

Los principales propósitos de la evaluación del aprendizaje del INEE son:

- Conocer el logro académico de los estudiantes a niveles estatal y nacional, así como los factores de contexto más importantes que explican las diferencias de los sectores estudiados.
- Contribuir al conocimiento de los alcances y limitaciones del Sistema Educativo Nacional y, con ello, promover el nivel de la calidad de la educación básica en nuestro país.
- Emitir juicios de valor contextualizados que sirvan para apoyar una toma de decisiones documentada.
- Complementar los procesos evaluativos existentes, desarrollados por otras instancias y programas nacionales e internacionales (como DGE y PISA).
- Aportar elementos para enriquecer la rendición de cuentas a que tiene derecho la sociedad mexicana.

Asimismo, el INEE ha definido que la evaluación del aprendizaje de gran escala que realice debe partir de los siguientes principios:

- Ser considerada como una evaluación externa del SEN del país.
- Ser de alta calidad y apegarse a normas y prácticas reconocidas internacionalmente.
- Proporcionar información que dé una imagen válida y confiable del SEN en su conjunto.
- Asegurar que las evaluaciones respeten el valor de equidad, en particular las relacionadas con las diferencias de género, de capital sociocultural y de grupos étnicos.

- Ser realizada de manera clara y transparente, con la participación colegiada de maestros y especialistas.
- Estar legitimada académica y socialmente.
- Proporcionar elementos que ayuden a mejorar la calidad del sistema educativo del país.

Marco de referencia de los EXCALE

Las bases documentales que soportan la nueva generación de pruebas nacionales se apoyan en el conocimiento acumulado durante más de 100 años de la psicometría, así como en la experiencia de aproximadamente 45 años de la aplicación de evaluaciones internacionales del logro educativo. Los trabajos que más influyeron en la elaboración de este marco de referencia² fueron los relacionados con: *a)* los fundamentos de las teorías Clásica de la medida y de la Respuesta al ítem, *b)* la experiencia y los marcos de referencia de las pruebas nacionales de otros países para evaluar la calidad educativa, especialmente los de España y Estados Unidos, y *c)* de las pruebas internacionales de vanguardia, como el Programa Internacional de Evaluación de Estudiantes (OCDE, 2003) y el Tercer estudio internacional de evaluación de ciencias y matemáticas (Schmidt *et al.*, 1997).

El marco teórico de los EXCALE está definido básicamente por los principios de las teorías Clásica de la medida (Nunnally y Bernstein, 1994) y de Respuestas al ítem (Hambleton, 1993). Respecto de la primera se adoptan los principios para la construcción y validación de pruebas criterioales alineadas al currículum; sobre la segunda, los principios básicos para la calibración y el escalamiento de pruebas educativas. De manera específica, la nueva generación de pruebas nacionales adopta los siguientes criterios teóricos-metodológicos:

- 1) ser de tipo criterial,
- 2) estar alineada al currículum,
- 3) tener un diseño matricial de reactivos,
- 4) ser de respuesta seleccionada (aunque en algunos casos también se utiliza la respuesta construida),
- 5) utilizar un escalamiento basado en la teoría de Respuestas al ítem,
- 6) definir niveles de logro para la interpretación de sus resultados, y
- 7) utilizar los parámetros de la teoría Clásica de la medida y de la Respuestas al ítem para evaluar la calidad de los EXCALE y aportar evidencias de su validez.

Los EXCALE son pruebas *criteriales* porque se diseñan para conocer con precisión el grado de dominio que el estudiante tiene sobre un conjunto de contenidos específicos. De tal modo, el referente para interpretar los resultados de los EXCALE es la cantidad y tipo de material que el estudiante maneja del universo de contenidos evaluados, o constructo de la prueba (Popham, 1990). Estas pruebas nos ayudarán a evaluar la posesión de conocimientos y habilidades escolares; a diferencia de las normativas que se utilizan con propósitos de ordenamiento y selección de individuos y donde se pone menor énfasis en los contenidos específicos que evalúan.

Por otro lado, los EXCALE son pruebas *alineadas al currículum* porque se elaboran con una metodología *ad hoc* para evaluar, con mucha precisión, los contenidos curriculares (llámense habilidades, conocimientos, competencias, etcétera) que se definen en los planes y programas de estudios nacionales (Contreras, 2000). Este alineamiento curricular implica, según Nitko (1994): identificar los resultados importantes pretendidos por el currículum, asociar las acciones de evaluación con los contenidos esenciales, definir el dominio curricular completo sobre el que se desarrollará el examen y precisar los resultados del aprendizaje que se establecen en el currículum oficial. México cuenta con un currículum nacional, libros de textos gratuitos y una formación magisterial más o menos uniforme, condiciones idóneas para este tipo de pruebas; a diferencia de otros países (como Estados Unidos) que para evaluar la educación a nivel nacional tienen que darse a la enorme tarea de generar estándares de contenido y ejecución en cada disciplina a fin de alinear con ellos el diseño de sus pruebas.

Son pruebas de tipo *matricial* porque están diseñadas para evaluar una gran cantidad de contenidos curriculares, sin someter a los estudiantes a largas jornadas de resolución de pruebas. Para ello se requiere construir un conjunto de reactivos que cubran el dominio completo de los contenidos curriculares a evaluar, para después dividirlo en subconjuntos y repartirlos entre los estudiantes, de tal manera que cada alumno conteste sólo algunos de ellos (Deng, Ferris y Hombo, 2003; Van der Linden, Veldkamp y Carlson, 2004). Con el modelo matricial se limita el número de reactivos que contesta cada estudiante (reduciendo con ello el tiempo de la evaluación), a la vez que permite cubrir la totalidad de contenidos curriculares seleccionados entre todos los sustentantes. Como consecuencia de lo anterior, la calificación individual del estudiante pierde precisión –en tanto

que su error de medida aumenta—; no así los resultados agregados a nivel de entidad federativa, modalidad educativa y estrato social, los que son el centro de interés del INEE (Gaviria, 2005).

Siendo los EXCALE pruebas de gran escala, sus reactivos son básicamente de *respuesta seleccionada*, con un diseño de opción múltiple. Todos los reactivos de este tipo contienen cuatro posibles opciones de respuesta, de las que una es correcta; no se utilizan opciones parcialmente correctas. Sin embargo, también hay reactivos de respuesta abierta o construida, en las pruebas de español, para evaluar la expresión de la lengua escrita.

La *calibración y el escalamiento* de las puntuaciones de los EXCALE se realizan de acuerdo con los principios y presupuestos de la teoría de Respuestas al ítem (Hambleton, 1993), específicamente utilizando el modelo de un parámetro, mejor conocido como modelo de Rasch (Wright, 1996; Linacre, 2005). Uno de los principios fundamentales de esta teoría es la “unidimensionalidad” de las escalas; es decir, se debe de comprobar que los reactivos de las pruebas corresponden a una sola dimensión, de lo contrario no es adecuado utilizar este modelo de escalamiento.

El *sistema de interpretación* de los EXCALE es uno de los elementos básicos de su validez; por ello los resultados de estas pruebas nacionales se interpretan de acuerdo con los estándares o niveles de logro que se definen para describir lo que los alumnos saben o son capaces de hacer con sus aprendizajes. El establecimiento de dichos niveles se realiza con base en el método de “Correspondencia de ítems”, particularmente con el más destacado de ellos, conocido como “marcador” (*bookmark*), descrito por Lewis, Mitzel, Green y Patz (1999). Como bien señaló Jornet (2005), la utilidad de la información que producen estas pruebas depende de la forma en que se den a conocer los resultados de aprendizaje del Sistema Educativo Nacional, la cual debe de permitir que se establezcan pautas precisas para la mejora del mismo.

Finalmente, con base en los principios de la teoría Clásica de la medida y de Respuesta al ítem, los EXCALE deberán ajustarse a rigurosos *estándares de calidad*, entre otros: *a)* la definición clara de su uso y cobertura, *b)* el uso de procedimientos rigurosos para su diseño y construcción, *c)* el uso de procedimientos estandarizados para su administración, *d)* la clara interpretación de resultados, y *e)* la exhibición de evidencias de validez y confiabilidad. La *validez* de los EXCALE se debe centrar, especialmente, en la premisa de que las puntuaciones de la prueba muestran qué tanto los

estudiantes saben y pueden hacer respecto al currículum nacional (Ruiz-Primo y Jornet, 2004).

Proceso de diseño, construcción, aplicación y validación de los EXCALE

Es importante mencionar que para poder diseñar pruebas nacionales de calidad se requiere seguir los lineamientos y estándares rigurosos para instrumentos de evaluación educativa (AERA, APA y NCME, 1999; Martínez y cols., 2000). Asimismo, es importante adoptar una metodología robusta que haya probado su validez para los propósitos que se persiguen; este es el caso del modelo de construcción de pruebas criterioales alineadas al currículum, donde un aspecto central es el trabajo colegiado de especialistas y docentes. Por lo anterior, el INEE adoptó para los EXCALE este modelo (véase Nitko, 1994), adaptándolo a las necesidades nacionales.

Así, el proceso quedó definido en siete fases y 16 etapas básicas. La tabla 1 muestra el proceso de diseño, construcción y validación de las pruebas nacionales del INEE. En ella se pueden apreciar, con detalle, los procedimientos y productos que se esperan de cada etapa, así como el personal externo al INEE que participa en cada una de ellas. Es importante señalar que, por lo general, los productos de cada etapa sirven como insumos de la siguiente, por lo que en el proceso de generación de este tipo de pruebas se considera, en parte, el de su validación (Contreras, 2000; Contreras, Backhoff y Larrazolo, 2003).

Como se puede apreciar en esta tabla y en la figura 1, a lo largo de todo el proceso participan diversos especialistas, grupos de asesores, comités de expertos y el personal técnico de la Dirección de Pruebas y Medición. Asimismo, en cada etapa del proceso se utilizan diversos procedimientos, entre los cuales destacan: 1) la documentación de procesos similares de construcción de pruebas de aprendizaje de gran escala, realizados por instituciones de reconocida calidad internacional; 2) la capacitación dirigida a los cinco comités de especialistas y docentes que participan en el proceso; 3) la elaboración y preparación de materiales para el trabajo de los cinco comités; 4) el trabajo colegiado, donde se toman decisiones de mayor importancia (tipo de diseño matricial); y 5) el trabajo individual por encargo, que requiere la experiencia y conocimientos de especialistas en un tema en particular (como diseño de la muestra poblacional).

TABLA 1

Proceso de diseño, construcción, aplicación y validación de los EXCALE

Etapas	Participantes externos*	Procedimientos	Productos
FASE I. PLANEACIÓN GENERAL**			
1. Diseño del Plan General de Evaluación	<ul style="list-style-type: none"> • Consejo técnico • Asesores en medición y validación 	<ol style="list-style-type: none"> 1. Documentación 2. Seminarios 3. Trabajo colegiado 4. Trabajos por encargo 	<ol style="list-style-type: none"> 1. Plan general de evaluación del aprendizaje, con el marco de referencia de los EXCALE 2. Manual general de procedimientos 3. Manual técnico para el diseño matricial 4. Manual técnico para el escalamiento de puntuaciones y niveles de competencias 5. Marco teórico de validación de los EXCALE
2. Diseño y elaboración de cuestionarios de contexto	<ul style="list-style-type: none"> • Especialistas en diseño de cuestionarios • Especialistas en evaluación del aprendizaje 		<ol style="list-style-type: none"> 6. Marco de referencia de los cuestionarios de contexto 7. Cuestionarios de contexto del alumno, docente y director
3. Diseño y desarrollo del sistema informático***	<ul style="list-style-type: none"> • Especialistas en bases de datos 		<ol style="list-style-type: none"> 8. Sistema de bases de datos de reactivos 9. Documento que describe la estructura y funcionamiento de la base de datos
FASE II. ESTRUCTURACIÓN DE LOS EXCALE			
4. Diseño de las pruebas	<ul style="list-style-type: none"> • Comités académicos (1 por prueba) 	<ol style="list-style-type: none"> 1. Documentación 2. Capacitación 3. Preparación de materiales 4. Trabajo colegiado 	<ol style="list-style-type: none"> 10. Manual técnico para el diseño de pruebas nacionales 11. Retícula curricular de cada prueba 12. Tabla de contenidos de cada prueba
5. Especificación de reactivos	<ul style="list-style-type: none"> • Comités elaboradores de especificaciones de reactivos (1 por prueba) 		<ol style="list-style-type: none"> 13. Manual técnico para la elaboración de especificaciones 14. Especificaciones de reactivos de cada prueba 15. Dos revisiones por especificación de reactivos
FASE III. CONSTRUCCIÓN DE REACTIVOS DE LOS EXCALE			
6. Elaboración de reactivos	<ul style="list-style-type: none"> • Comités constructores de reactivos (1 por prueba) 	<ol style="list-style-type: none"> 1. Documentación 2. Capacitación 3. Trabajo individual y colegiado 	<ol style="list-style-type: none"> 16. Manual técnico para la construcción de reactivos 17. Tres reactivos por especificación 18. Dos revisiones por reactivo construido
7. Validación de reactivos	<ul style="list-style-type: none"> • Comités de validación y sesgo (1 por prueba) 		<ol style="list-style-type: none"> 19. Manual técnico para la validación de reactivos 20. Dos reportes de validación por cada reactivo
8. Piloteo de reactivos y cuestionarios de contexto	<ul style="list-style-type: none"> • Algunas coordinaciones estatales de evaluación 	<ol style="list-style-type: none"> 1. Muestreo para el piloteo 2. Predicción de reactivos e impresión de cuadernillos 3. Capacitación 	<ol style="list-style-type: none"> 21. Manual técnico para el piloteo de reactivos 22. Muestra poblacional 23. Manual técnico para la edición de reactivos 24. Cuadernillos de prueba y cuestionarios de contexto impresos 25. Base de datos con resultados de pilotaje

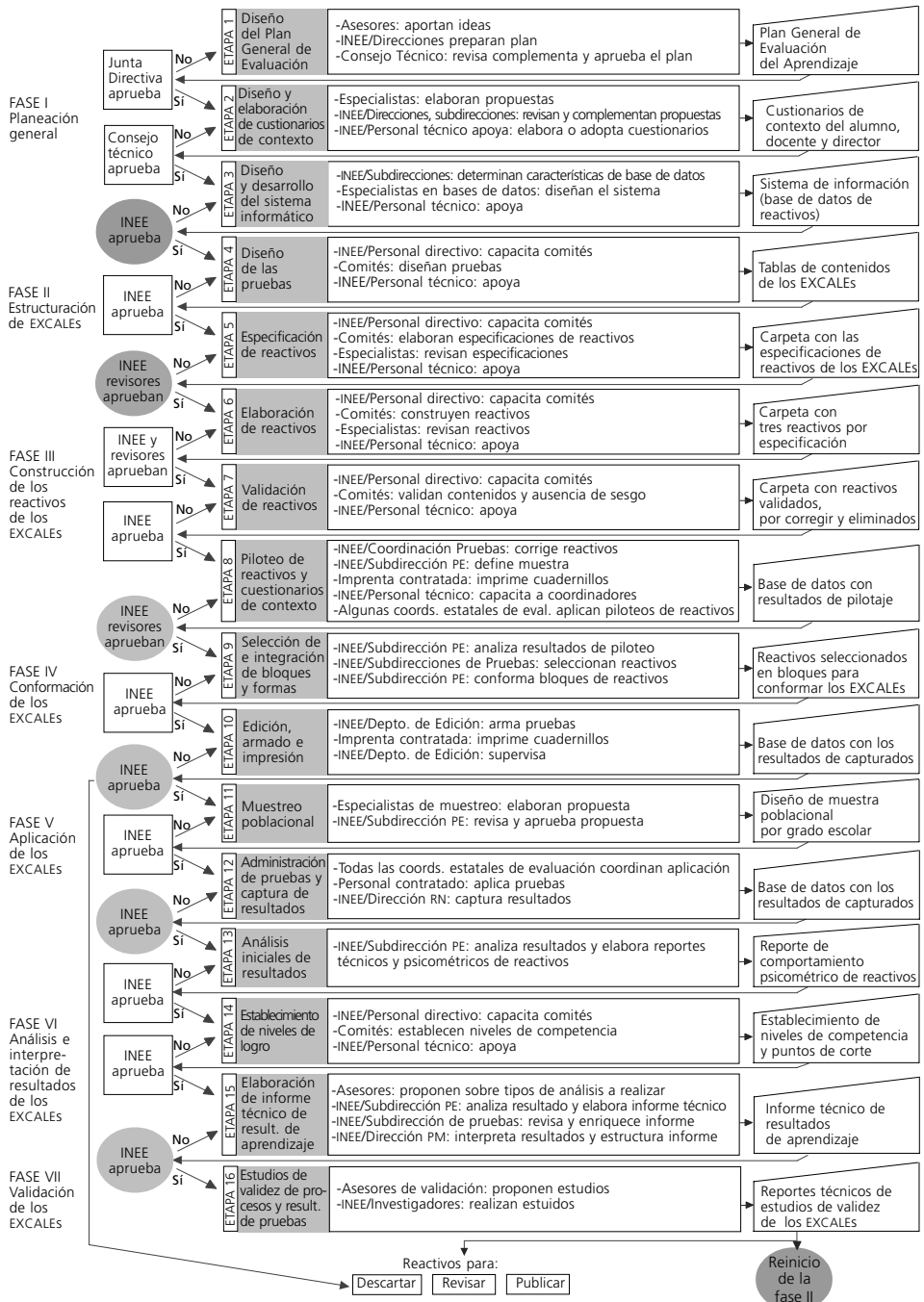
Etapas	Participantes externos*	Procedimientos	Productos
FASE IV. CONFORMACIÓN DE LOS EXCALE			
9. Selección de reactivos e integración de bloques y formas	• Asesores en medición	1. Documentación 2. Análisis de reactivos	26. Manual técnico para el análisis psicométrico de reactivos 27. Informe de estimadores estadíst. de react. 28. Bloques de reactivos 29. Estructura de formas (combinación de bloques)
10. Edición, armado e impresión	• Imprenta contratada	1. Edición de cuadernillos y cuestionarios de contexto 2. Impresión de cuadernillos de lectura óptica	30. Manual técnico para la edición de reactivos y cuadernillos 31. Cuestionarios de contexto impresos 32. Cuadernillos de los EXCALE impresos
FASE V. APLICACIÓN DE LOS EXCALE			
11. Muestreo poblacional	• Especialistas en muestreo	1. Documentación 2. Trabajo colegiado 3. Trabajo por contrato	33. Manual técnico para seleccionar muestras poblacionales 34. Marco muestral actualizado 35. Diseño y muestra poblacional
12. Administración de las pruebas y captura de resultados	• Todas las coordinaciones estatales de evaluación • Personal contratado	1. Documentación 2. Distribución de cuadernillos en las entidades	36. Manual técnico para la administración de cuadernillos y cuestionarios de contexto 37. Paquetes de cuadernillos distribuidos en las entidades federativas 38. Base de datos con resultados de la evaluación capturados
FASE VI. ANÁLISIS E INTERPRETACIÓN DE RESULTADOS DE LOS EXCALE			
13. Análisis iniciales de resultados	• Asesores en medición	1. Seminarios 2. Análisis estadísticos de resultados	39. Manual técnico sobre análisis de reactivos 40. Informe técnico sobre comportamiento psicométrico de reactivos
14. Establecimiento de niveles de logro	• Comités de niveles de logro (dos por prueba)	1. Documentación 2. Capacitación 3. Trabajo colegiado	41. Manual técnico sobre establecimiento de niveles de logro 42. Documento donde se consignen los niveles de logro y puntos de corte de cada prueba
15. Elaboración de informe técnico de resultados de aprendizaje	• Asesores en medición y validación	1. Documentación 2. Capacitación 3. Trabajo colegiado	43. Manual técnico sobre informe de resultados de aprendizaje 44. Informe técnico sobre resultados de aprendizaje asociados con variables de contexto
FASE VII. VALIDACIÓN DE LOS EXCALE			
16. Estudios de validez de procesos y resultados de las pruebas	• Asesores en validación	1. Documentación 2. Investigación	45. Marco de referencia de los estudios de validez 46. Marco de referencia de cada prueba 47. Reportes técnicos de estudios de validez 48. Publicaciones sobre la validez de las pruebas

* Personal de la Dirección de Pruebas y Medición interviene en todas las etapas del proceso.

** Fase general para la nueva generación de las pruebas EXCALE.

*** A partir de la etapa 4 el sistema informático se irá alimentando con la información que se produzca a lo largo de todo el proceso.

FIGURA 1. Diagrama del proceso de diseño y desarrollo de las pruebas EXCALE



A continuación se describen brevemente cada una de las siete fases y 16 etapas del proceso de diseño, construcción, aplicación y validación de esta nueva generación de pruebas.

Fase I: Plan general

La primera fase del proceso tiene el propósito principal de establecer un plan de pruebas a largo plazo, por lo que resulta de particular importancia la participación del Consejo Técnico y de los asesores externos en medición y validación; encargados de definir los propósitos, principios y referentes conceptuales de los que dependerán las fases restantes del proceso. Esta fase consta de tres etapas:

En la primera se diseña el Plan general de evaluación del aprendizaje, donde se plasman: el marco de referencia de los EXCALE; el proceso de diseño, construcción, aplicación y validación de pruebas; el diseño matricial; y el modelo de niveles de logro. Asimismo, se elabora el marco de teórico del aprendizaje escolar y el programa para validar las interpretaciones de las nuevas pruebas nacionales (INEE, 2004b).

En la segunda etapa se diseñan y elaboran los cuestionarios de contexto –dirigidos a estudiantes, docentes y directores de escuela– que se aplicarán junto con las pruebas para explicar los resultados de aprendizaje que se obtengan.

Finalmente, en la tercera etapa, se diseña y desarrolla un sistema automatizado de información para alojar, mantener y manejar la base de datos (estructura curricular de la asignatura, tabla de contenidos de la prueba, especificaciones y plantillas de reactivos, resultados de la validez, parámetros del piloteo de reactivos, entre otros) relativa a las distintas pruebas que se vayan generando.

Fase II: Estructuración de los EXCALE

A partir de esta segunda fase todas las etapas del proceso son específicas para cada prueba. Esta fase tiene el propósito de diseñar y justificar la estructura del examen y, por otra parte, elaborar las especificaciones de todos los reactivos que conformarán la prueba. En ella participan especialistas en currículum, en la enseñanza de la disciplina, autores de libros de texto, representantes de asociaciones, así como docentes en ejercicio pertenecientes a los distintos estratos y modalidades educativos.

En esta fase se establecen la cuarta y quinta etapas del proceso. En la cuarta etapa el Comité académico de cada prueba, formado por aproxi-

madamente una decena de especialistas, realiza un análisis curricular exhaustivo de la asignatura y grado correspondientes, a fin de generar una retícula (tabla de doble entrada) de la asignatura donde se explicita la estructura curricular y los contenidos que son esenciales y muy importantes de evaluar. A partir de la retícula, se elabora la tabla de contenidos de la prueba y la justificación de los contenidos curriculares que la conforman. En esta tabla se explicitan las áreas temáticas o componentes, los temas y subtemas que derivan en los contenidos curriculares y habilidades intelectuales a evaluar, así como la forma y número de reactivos con que serán evaluados.

En la quinta etapa, el Comité elaborador de especificaciones de reactivos de cada prueba, formado por una decena de especialistas y docentes, define y describe en forma detallada las características que debe tener cada reactivo. Es decir, se precisa el contenido a evaluar, su ubicación en el currículum, su importancia, la habilidad intelectual requerida al alumno y el formato del reactivo mismo (características de forma, fondo, redacción, etcétera), de tal manera que esta descripción sirva como guía o molde para que se construyan reactivos similares y, hasta donde sea posible, equivalentes. El número de especificaciones para cada prueba es variable, pues depende de la extensión curricular de la asignatura.

Fase III: Construcción de reactivos de los EXCALE

La tercera fase del proceso corresponde a la elaboración de los reactivos, su validación y piloteo. Su propósito es producir reactivos de alta calidad para la conformación de la prueba. Dentro de esta fase se ubican la sexta, séptima y octava etapas del proceso.

En la sexta etapa, miembros del Comité constructor de reactivos de cada prueba, constituido aproximadamente por una decena de autores de libros de texto y docentes en ejercicio formulan, de manera independiente, tres reactivos para cada especificación, que serán revisados por otros especialistas según el manual de revisión de reactivos, haciendo énfasis principalmente en la congruencia reactivo-especificación. El proceso se repite hasta que los reactivos quedan terminados a satisfacción de los revisores.

Los reactivos resultantes se pre-editan, se alojan en el sistema informático, y pasan al Comité de validez y sesgo con el cual se inicia la séptima etapa del proceso. Cabe resaltar que dicho Comité está constituido por docentes

en ejercicio de las 32 entidades federativas del país y representantes de los diversos estratos y modalidades educativos. Este Comité revisa cada reactivo en relación con su contenido (pertinencia curricular, grado de dificultad), redacción (lenguaje utilizado, construcción sintáctica) y sesgos cultural y de género (características de contenido y redacción que favorezcan o perjudiquen a algún grupo social). Cada reactivo es revisado independientemente por dos profesores y, en caso de discrepancia, por un grupo de ocho docentes quienes deben llegar a consenso. En su caso, se proponen recomendaciones para mejorar los reactivos o se justifica su eliminación.

En la octava etapa se realiza un estudio piloto con una muestra intencional de aproximadamente cinco mil estudiantes por prueba, con el fin de conocer el comportamiento psicométrico de los reactivos y detectar los problemas que enfrentan los alumnos al responderlos.³ Este estudio piloto ensaya las condiciones reales de aplicación, mismas que incluyen la capacitación a coordinadores y responsables de evaluación, la aplicación de encuestas a alumnos, docentes y directores, así como la logística completa de los estudios evaluativos. Esta etapa culmina con la lectura de las respuestas de los estudiantes y la conformación de una base de datos con los resultados obtenidos.

Fase IV: Conformación de los EXCALE

El propósito de esta fase es editar los cuadernillos de cada una de las pruebas nacionales con la información recabada en las dos etapas anteriores; se seleccionan los reactivos que conformarán las pruebas nacionales, se agrupan en bloques y se definen las diversas formas de cada una de las pruebas.

Para llevar a cabo dicha selección se efectúa un análisis del comportamiento psicométrico de los reactivos (en diversos grupos de estudiantes) y se consideran las observaciones efectuadas por los docentes en cuanto a su validez de contenido y ausencia de sesgo. Este análisis se realiza considerando los estándares de las pruebas criteriosales, es decir, dándole mayor preferencia al aspecto de contenido que al comportamiento psicométrico.

En la décima etapa, se forman distintos bloques de reactivos y con ellos diversas formas. Los bloques se conforman con una pequeña cantidad de reactivos que, en conjunto, puedan ser contestados en 15 minutos y que compartan ciertas características que los hagan equivalentes en cuanto a

su contenido, nivel de dificultad y varianza, así como en su extensión. Una vez conformados los bloques de reactivos, éstos se combinan con el fin de armar distintas formas, de tal manera que todos los bloques se encuentren, hasta donde sea posible, en igual proporción y distribución. El número de formas puede variar de una prueba a otra. Finalmente, a cada forma se le agrega el cuestionario de contexto del alumno para mandar a impresión los cuadernillos resultantes y con ello concluir esta fase.

Fase V: Aplicación de los EXCALE

La quinta fase del proceso tiene como propósito principal administrar las pruebas considerando una muestra nacional de estudiantes, así como los cuestionarios de contexto que serán aplicados a alumnos, docentes y directores. Asimismo se desarrolla la captura de los resultados en una base de datos para su posterior análisis.

Durante esta fase se diseña la muestra poblacional de alumnos de conformidad con los estudios evaluativos que se planeen realizar; se capacita al personal encargado de la coordinación de la aplicación en las entidades federativas; se distribuyen los cuadernillos y, finalmente, se aplican las pruebas nacionales a las muestras poblacionales seleccionadas.

El diseño de estas muestras se realiza en la etapa décimo primera; acción que implica la selección de escuelas y estudiantes de las 32 entidades federativas del país. El diseño depende de muchos factores, entre los que se encuentran: el marco muestral disponible, los propósitos del(los) estudio(s) a realizarse, la representatividad de las subpoblaciones que se deseen evaluar (rurales, urbanas, privadas, etcétera), el grado de confianza de la muestra, y el tipo de diseño matricial de las pruebas. Una vez que se diseña la muestra aleatoriamente, ésta se valida con la información que proporcionan las escuelas seleccionadas.

En la décimo segunda etapa se realiza el proceso de capacitación a los coordinadores y responsables de la aplicación (cerca de 5000), se distribuyen los cuadernillos a las 32 entidades federativas, se administran las pruebas a los estudiantes seleccionados, se aplican los cuestionarios de contexto a docentes y directores de escuelas de los dos niveles educativos y, por último, se leen y se capturan los resultados de las pruebas y cuestionarios en una base de datos. Es pertinente hacer notar que, debido al arreglo matricial de la aplicación, la base de datos resultante tiene particularidades que causan que algunos análisis estadísticos tengan que realizarse de manera dis-

tinta a la usual, y que otros tengan que sustituirse por análisis *ad hoc* para este tipo de pruebas.

Para el caso de los reactivos de respuesta abierta donde sólo se utiliza una muestra nacional sin representatividad estatal, como es el caso de la sección de redacción de la prueba de español, un par de jueces califican conjuntamente cada reactivo con base en una rúbrica o protocolo de calificación para que, posteriormente, se agreguen los resultados a la base de datos correspondiente. Esta etapa del proceso es muy delicada, pues en ella se presentan una diversidad de problemas que fácilmente podrían invalidar los resultados de las evaluaciones. Entre otras acciones para fortalecer el procedimiento, se lleva a cabo un periodo de estandarización de criterio de los jueces, y se hacen estudios de confiabilidad entre ellos.

Fase VI: Análisis e interpretación de resultados de los EXCALE

Esta fase consta de tres etapas y tiene como propósito final la elaboración de los informes técnicos sobre los resultados de las pruebas y los cuestionarios de contexto; elementos que serán el sustento de los reportes anuales relacionados con el aprendizaje que publique el INEE. Para lograr esta meta, es necesario realizar un análisis inicial de los resultados con el fin de establecer los niveles de logro de los estudiantes e incluirlos en los reportes técnicos correspondientes.

Así, en la etapa décimo tercera, se realizan los primeros análisis estadísticos, tanto de estudiantes como de reactivos, poniendo especial énfasis en el comportamiento psicométrico de estos últimos.

En la etapa décimo cuarta, el Comité de establecimiento de niveles de logro de cada prueba, compuesto por autoridades educativas, especialistas en currículum, autores de libros de texto y docentes en ejercicio, define las categorías y niveles de logro de los estudiantes de acuerdo con: 1) la ejecución que en “teoría” se esperaría de ellos y, 2) los resultados reales en la prueba respectiva. Con esta información se definen los puntos de corte de cada nivel de logro, es decir, las puntuaciones mínima y máxima que, en la escala de puntuaciones, corresponden a cada nivel.

Finalmente, en la etapa décimo quinta, se elaboran los informes técnicos sobre los resultados de aprendizaje de los estudiantes evaluados, considerando para ello las variables de contexto y de oportunidades de aprendizaje aportadas en los cuestionarios de contexto aplicados a alumnos, docentes y directores.

De especial interés para el INEE es conocer el logro del aprendizaje en las distintas subpoblaciones muestreadas (entidades federativas y estratos y modalidades educativos), así como los contenidos que los estudiantes dominan del currículum nacional. No está por demás subrayar que la información generada en esta etapa nutre significativamente la publicación anual del INEE: *La calidad de la educación en México*.

Fase VII: Validación de los EXCALE

Por último, esta fase consta de una sola etapa y tiene como meta aportar información diversa sobre la validez de las pruebas y de las interpretaciones que se deriven de su uso.

Aunque para fines prácticos la etapa décimo sexta se ubica al final del proceso de construcción de las pruebas, en realidad inicia desde el momento en que termina el diseño (cuarta etapa). Podemos ver dos momentos de los estudios de validez de los EXCALE: *a)* durante el proceso mismo de su construcción y *b)* posterior a su terminación.

En el primer caso los estudios van dirigidos a evaluar la calidad y congruencia de cada una de las etapas del proceso, y tiene como finalidad verificar el proceso mismo, así como generar información tendiente a corregir los problemas detectados mientras se construye la prueba.

En el segundo caso los estudios tienen el propósito de aportar evidencias de la veracidad y limitaciones de las interpretaciones que se generen como resultado del uso de una prueba. Su finalidad es doble: por un lado, legitimar académicamente las evaluaciones generadas por el INEE y, por el otro, iniciar un proceso de mejora de las pruebas con base en la información y documentación de estos estudios. El proceso de validación de una prueba nunca termina; sólo se van acumulando evidencias de la veracidad o falsedad de sus interpretaciones.

Conclusiones

Queda claro que para el INEE el objetivo fundamental de la evaluación del aprendizaje es conocer el rendimiento académico de los estudiantes a niveles estatal y nacional, así como las oportunidades de aprendizaje y los factores de contexto que lo explican. Lo anterior, con el fin de emitir juicios de valor que apoyen la toma de decisiones documentada y contribuir a realizar una rendición de cuentas a la sociedad mexicana sobre el estado que guarda la educación nacional (Poder Ejecutivo, 2002).

Asimismo, se espera que los resultados de las evaluaciones que realice el INEE impactarán los siguientes ámbitos educativos: políticas nacionales y estatales, currículum nacional, programas de estudio y libros de texto, administración y gestión escolares, formación y actualización docentes, y opinión de la sociedad (INEE, 2004b). Para lograr este impacto el INEE debe contar, por un lado, con instrumentos y procedimientos evaluativos que sean teórica y metodológicamente sólidos; y, por el otro, que los resultados de sus evaluaciones que se hagan públicos sean válidos y confiables, de tal manera que tanto docentes como autoridades educativas puedan tomar decisiones documentadas para la mejora de la educación al igual que tanto académicos como la sociedad misma puedan ejercer presión a las autoridades correspondientes para que las medidas de mejora educativa se tomen de manera informada.

Es importante subrayar dos características de la metodología empleada para diseñar, construir, aplicar y validar los EXCALE: *a)* el trabajo colegiado, donde intervienen una gran cantidad de especialistas y docentes en ejercicio, y *b)* la documentación que se genera a lo largo de todo del proceso, que deja evidencia de su validez. En cuanto al trabajo colegiado, hay que destacar la participación de cinco comités, cada uno conformado aproximadamente por diez personas; la asesoría de dos grupos de expertos, integrado cada uno por tres especialistas; el trabajo por encargo de alrededor de cinco profesionales, y el trabajo especializado de cerca de 25 técnicos de la Dirección de Pruebas y Medición. Es decir, para el diseño, construcción y validación de cada EXCALE se requiere de casi un centenar de especialistas y docentes. Esto sin considerar las aproximadamente 900 personas que interviene en el estudio piloto y las cerca de cinco mil que lo hacen en la aplicación nacional.

Sobre la documentación que se genera con esta metodología, debemos destacar tres aspectos. Primero, en su gran mayoría, los productos se van elaborando en forma secuenciada y cada uno de ellos sirve de insumo para el siguiente paso. Segundo, los productos generados en cada etapa son revisados y analizados por el siguiente comité, y aunque éste no puede modificar lo hecho por el grupo anterior, sí debe señalar y, en su caso, justificar las observaciones que se hagan al respecto, de tal manera que se tenga evidencia de los problemas que se deben resolver en el futuro para mejorar el instrumento. Tercero, la documentación que se produce al final del proceso es abundante y rica en información; cerca de

48 productos se elaboran para cada prueba, de los cuales aproximadamente la mitad son comunes a todas ellas (como los manuales técnicos) y la mitad restante son específicos de cada examen (como las tablas de contenidos).

Con esta metodología creemos que es posible alcanzar dos de las grandes metas que se trazó el INEE respecto de la evaluación del aprendizaje:

- 1) Conocer el logro académico de los estudiantes a niveles estatal y nacional, así como los factores de contexto más importantes que lo determinan.
- 2) Contribuir a conocer los alcances y limitaciones del Sistema Educativo Nacional y, con ello, el nivel de la calidad de la educación básica en nuestro país.

Estas metas se pueden lograr gracias a las bondades del modelo de los exámenes alineados al currículum. Ése es el caso de los EXCALE, donde el currículum es la base sobre la que se construye el examen, y las decisiones sobre qué evaluar y cómo hacerlo están determinadas por los resultados de aprendizaje que se establecen en dicho currículum. El punto nodal de este modelo es poder garantizar que el conjunto de reactivos que llamamos EXCALE representa al universo de contenido que llamamos currículum nacional; desde luego, la clave para garantizarlo es el juicio humano, el cual está presente a lo largo de todo el proceso de diseño y validación de estos exámenes, tal como lo propone Nitko (1994).

Para finalizar debemos decir que este documento, que explica de manera general el modelo de la nueva generación de pruebas, representa un primer paso para dar a conocer al público los aspectos más relevantes de la metodología empleada. Sin embargo, este texto se deberá tomar con las reservas necesarias pues, con toda seguridad, el modelo de evaluación aquí expuesto se irá modificando y fortaleciendo conforme se vaya recibiendo información de los resultados de su operación.

El INEE, siendo coherente con sus principios, tiene el compromiso de informar a la sociedad de sus métodos y resultados, razón por la cual se publica este escrito. Lo anterior toma una gran relevancia si consideramos que la evaluación debe ser un proceso documentado y abierto; y no como en el pasado que las evaluaciones nacionales fueron consideradas “cajas negras” a las cuales muy pocas personas tuvieron acceso.

Notas

¹ Los grados que se planean evaluar son los terminales de la educación básica y media superior (3º de preescolar, 6º de primaria, 3º de secundaria y 3º de bachillerato), así como el 3º de primaria por ser el grado intermedio.

² Es común que en las pruebas de logro educativo se hable del “marco de referencia” en vez

del marco teórico, ya que el primero es más comprensivo e incluye, además de posturas teóricas, las características de las pruebas, el modelo metodológico de construcción de pruebas y el marco conceptual de los dominios que evalúa.

³ Para lo cual se requiere el trabajo de cerca de 900 aplicadores.

Referencias bibliográficas

- American Educational Research Association (AERA), American Psychological Association (APA) y National Council on Measurement in Education (NCME) (1999). *Standards for educational and psychological testing*, Washington: American Psychological Association.
- Backhoff, E. (2005). “La comparación entre entidades: alcances y limitaciones de los rankings”, en *Memorias de las jornadas de evaluación educativa*, México: INEE.
- Backhoff, E. y Martínez-Rizo, F. (2004). “Resultados de las pruebas de estándares nacionales 2003: elementos para la comparación entre entidades”, en *Memoria 2004, sexto foro de evaluación educativa*, México: CENEVAL.
- Contreras, L. A. (2000). *Desarrollo y pilotaje de un examen de español para la educación primaria en Baja California*, tesis para obtener el grado de maestro en Ciencias educativas, México: Universidad Autónoma de Baja California.
- Contreras, L. A.; Backhoff, E. y Larrazolo, N. (2003). *Curso taller para la elaboración de exámenes criterios: manual para el Comité diseñador del examen*, documento mimeografiado, Ensenada: Instituto de Investigación y Desarrollo Educativo-UABC.
- Deng, H.; Ferris, J. y Hombo, C. (2003). *A vertical scheme of building the naep booklets*, documento presentado en la Reunión anual del NCME, Chicago.
- DGE (2000). *Balance de las acciones emprendidas entre diciembre de 1994 y octubre de 2000*, documento recuperado el 24/05/2005 en: <http://www.sep.gob.mx/work/apps/site/dge/index.htm>
- DGE (2004). *La evaluación en la Secretaría de Educación Pública*, documento mimeografiado, México: SEP-DGE.
- Gaviria, J.L. (2005). *Propuesta de diseño matricial para las pruebas de español y matemáticas del programa de pruebas nacionales del INEE*, documento mimeografiado, México: INEE.
- Hambleton, R.K. (1993). “Principles and selected applications of item response theory”, en Linn (ed.), *Educational measurement* (3rd ed.), Nueva York: MacMillan Publishing Co, pp. 147-200.
- INEE (2003). *La calidad de la educación básica en México*, México: INEE.
- INEE (2004a). *La calidad de la educación básica en México: resultados de evaluación educativa 2004*, México: INEE.
- INEE (2004b). *Plan general de evaluación del aprendizaje*, documento mimeografiado, México: INEE.
- Jornet, J. (2005). *El modelo de determinación de estándares de los Exámenes de la Calidad y Logro Educativos (EXCALE) del INEE de México*, México: INEE.

- Lewis, D. M.; Mitzel, H. C.; Green, D. R. y Patz, R. J. (1999). *The bookmark standard setting procedure*, Monterey, CA: McGraw-Hill.
- Linacre, J. M. (2005). *Winsteps rasch measurement computer program*, Chicago: Winsteps.com.
- Marínez-Rizo, F. (2004). "Comparabilidad de los resultados de las evaluaciones", en *Memorias de las jornadas de evaluación educativa*, México: INEE.
- Martínez-Rizo, F. et al. (2000). *Estándares de calidad para instrumentos de evaluación educativa*, México: CENEVAL.
- Nitko, A. (1994, julio). *A model for developing curriculum-driven criterion-referenced and norm-referenced national examinations for certification and selection of students*, ponencia presentada en la Conferencia internacional sobre evaluación y medición educativas, de la Asociación para el Estudio de la Evaluación Educativa en Sudáfrica (ASSESA).
- Nunnally, J. C. y Bernstein, I. H. (1994). *Psychometric theory*, Nueva York: Mc Graw-Hill.
- OCDE (2003). *The PISA 2003, Assessment Framework: Mathematics, Reading, Science and Problem Solving*, documento mimeografiado. s/l: Organización para la Cooperación y Desarrollo Económicos.
- Poder Ejecutivo (2002). "Decreto de creación del Instituto Nacional para la Evaluación de la Educación", *Diario Oficial*, 08/08/2002, México, D. F.
- Popham, J. (1990). *Modern educational measurement: A practitioner's perspective*, Englewood Cliffs, N J: Prentice-Hall.
- Ruiz-Primo, M. A., Jornet, J. (2004). *Acerca de la validez de los Exámenes de la Calidad y el Logro Educativos (EXCALE)*, México: INEE.
- SEP (2001). *Programa Nacional de Educación 2001-2006*, México: SEP.
- Schmidt, W. et al. (1997). *Many visions, many aims*, vol. 2: *A cross-national investigation of curricular intensions in school science*, Dordrecht, Holanda: Kluwe Academic Publishers.
- Van der Linden, W.; Veldkamp, B. y Carlson, J. (2004). "Optimizing balanced incomplete block designs for educational assessments", *Applied Psychological Measurement*, 28(5), 317-331.
- Velázquez, V. (2000). "Hacia una cultura de la evaluación", en SEP, *Memorias del quehacer educativo 1995-2000*, México: SEP.
- Wright, B. D. (1996). "Local dependency, correlations and principal components", *Rasch Measurement Transactions*, 10, 3, 509-511.

Artículo recibido: 4 de agosto de 2005

Dictamen: 9 de noviembre de 2005

Segunda versión: 25 de noviembre de 2005

Aceptado: 19 de enero de 2006