

Marco de Trabajo de la Evaluación Nacional 2004

1. El aporte de las evaluaciones de sistema

Las evaluaciones de sistema cobran una gran importancia en la región latinoamericana y comienzan a ser desarrolladas junto con el diseño e implementación de las reformas educativas de los años 80 y 90. En gran parte, esto responde a la necesidad de contar con sistemas nacionales de evaluación que puedan ofrecer información sobre el sistema educativo y sus componentes, con la finalidad de colocar la calidad de la educación como tema central y prioritario de la agenda pública.^{1,2}

Existe consenso en que el logro de aprendizajes de los estudiantes es uno de los principales indicadores de la calidad del sistema educativo. En ese sentido, el sistema de medición de la calidad busca generar información sobre dichos logros y sobre un conjunto de factores escolares y extraescolares que inciden en los rendimientos de los estudiantes para poder explicarlos desde una perspectiva de equidad.

La evaluación en el campo educativo, implica un conjunto de procesos, que tienen como propósito obtener información sobre el aprendizaje de los estudiantes. Desde una perspectiva macro o de sistema, esta información está referida a las metas educativas del sector, lo que permite contrastar, a escala nacional, lo alcanzado por los estudiantes (currículo logrado) frente a lo prescrito en el currículo oficial (currículo intencional), en cambio, la perspectiva de evaluación micro o de aula permite obtener información sobre el desarrollo de los procesos formativos de cada uno de los alumnos del aula y así ajustar dichos procesos a las necesidades particulares de cada grupo de aula.³ Desde ambas perspectivas, se obtiene información para retroalimentar procesos y tomar decisiones pero a diferente nivel, cada una de ellas tiene su especificidad, sus propios instrumentos y procedimientos. Por ello, se puede afirmar que la evaluación de aula y la evaluación de sistema son diferentes debido a sus propias características y alcances, pero también complementarias porque ambas buscan información para retroalimentar, comparar y mejorar los aprendizajes de los estudiantes y por ende la calidad del servicio educativo.

2. Las evaluaciones de rendimiento en el Perú

La Unidad de Medición de Calidad Educativa (UMC) del Ministerio de Educación ha realizado Evaluaciones Nacionales (EN) en 1996, 1998 y 2001. En 1996 se administraron pruebas de comunicación y matemática⁴ en cuarto grado de primaria a escala nacional; esta muestra incluyó solamente estudiantes de centros educativos polidocentes completos de zonas urbanas y rurales (los centros educativos multigrado y unidocentes no fueron incluidos).

En 1998 se administraron pruebas a escala nacional en zonas urbanas en cuarto grado de primaria (matemática, comunicación, ciencia y ambiente y personal-social), sexto de primaria (matemática, lenguaje, ciencias naturales y ciencias sociales) y cuarto y quinto de secundaria (lenguaje y matemática) usando como base para el desarrollo de las pruebas el currículo vigente para cada grado. El marco de referencia de las pruebas de

¹ De Moura (1997)

² En el Perú, los resultados de las recientes evaluaciones a escala de sistema han proporcionado información sobre los aprendizajes de los estudiantes que han revelado, entre otras cosas, la necesidad de dar atención prioritaria al sector, mejorando la calidad del servicio educativo que proporciona el Estado. Esta información ha sido un elemento decisivo en la declaración de la emergencia educativa (2004-2007).

³ Benavides, *et al.* (2002)

⁴ Los nombres de las materias o áreas curriculares han variado en los diversos diseños y estructuras curriculares oficiales de los niveles primaria y secundaria, para efectos de facilitar la comprensión del presente documento se emplearán los términos genéricos: "matemática" y "comunicación".

las dos EN mencionadas ha sido normativo⁵, y como tal se utilizó para identificar grupos relativamente rezagados en rendimiento a nivel nacional (por ejemplo departamentos del Perú en los que el rendimiento se ubica consistentemente entre los más bajos a nivel nacional; y estudiantes de centros educativos públicos en relación a los privados).

La Evaluación Nacional de 2001⁶ tuvo importantes diferencias con las dos anteriores. En primer lugar, se abandonó el modelo de normas, ya que brindaba poca información sobre cuánto saben los estudiantes, para adoptar un modelo referido a criterios, que permitió establecer cuánto sabían los estudiantes peruanos respecto de algunas competencias del currículo. En segundo lugar, se utilizaron formatos de pregunta de opción múltiple pero también de respuesta abierta corta, de respuesta abierta extendida, de comunicación oral y escrita y de solución de problemas con material concreto. En tercer lugar, se incorporó a la muestra a estudiantes de zonas rurales y bilingües, adaptando para ello tanto las pruebas como los procedimientos de administración. Los resultados de la EN 2001 mostraron en general rendimientos por debajo de lo esperado dados los currículos vigentes. Estos resultados han sido reportados en una serie de documentos (ver página web de la UMC: www.minedu.gob.pe/mediciondelacalidad).

Por otro lado, el Perú ha participado en dos evaluaciones internacionales. En 1997 Perú participó junto con 11 países de la región en el Primer Estudio Regional Comparativo del Laboratorio Latinoamericano de Evaluación de la Calidad de la Educación (LLECE) de la UNESCO en el que se evaluó a estudiantes de tercero y cuarto grado de primaria en las áreas de matemática y comunicación. En segundo lugar, Perú ha participado en la evaluación del Programa Internacional para la Evaluación de Estudiantes (PISA) de la OCDE⁷, orientada a evaluar las habilidades en comprensión de lectura, matemática y ciencia de estudiantes de 15 años de secundaria. Esta evaluación ha sido especialmente útil para mostrar los estándares de los países desarrollados, es decir, qué es lo que esperan éstos respecto de los aprendizajes de los estudiantes de dicha edad. Estos exigentes estándares deberían ser considerados gradualmente y posiblemente incorporados como uno de los insumos a tener en cuenta en la definición de políticas educativas futuras en nuestro país.

Respecto a las características de la población evaluada, en todos los casos, las evaluaciones realizadas a escala de sistema tanto nacionales como internacionales han sido diseñadas a través de muestras probabilísticas y representativas de los estudiantes a nivel nacional y de los dominios de estudio correspondientes. Cabe mencionar que dichas evaluaciones tienen un propósito esencialmente formativo, no se han concebido para promover o sancionar a docentes o centros educativos, es decir, no tienen "consecuencias" en términos de estímulos ni de sanciones para los participantes en ellas.

⁵ El modelo de normas busca mostrar la posición relativa entre diferentes grupos, pero no ofrece información sobre el desempeño de los mismos.

⁶ En adelante, EN 2001

⁷ Organización para la Cooperación y el Desarrollo Económico.

El cuadro que se presenta a continuación resume las principales características del diseño de las evaluaciones realizadas por UMC:

Evaluaciones	Nombre	Modelo	Grados	Área	Población	Estratos de la Muestra *
Nacional	Crecer 96	Normas	4° Primaria	Comunicación Matemática	CE polidocentes completos	Urbanos y rurales Departamentos**
	Crecer 98	Normas	4° Primaria 6° Primaria 4° Secundaria 5° Secundaria	Comunicación Matemática Ciencia y ambiente (Primaria) Personal social (Primaria)	CE polidocentes completos urbanos	Departamentos***
	EN 2001	Criterios	4° Primaria 6° Primaria 4° Secundaria	Comunicación Matemática	CE del país	Primaria: - CE polidocentes completos y multigrados (incluye unidocentes) - EBI quechua (collao) - EBI aimara Secundaria: - Departamentos ***
Internacional	LLECE (1997)	Normas	3° Primaria 4° Primaria	Comunicación Matemática	CE del país	Megaciudad Urbana Rural
	PISA (2001)	Criterios	Secundaria (estudiantes de 15 años)	Alfabetización lectora Alfabetización matemática Alfabetización científica	CE del país ****	

* Todas las evaluaciones contemplaron niveles de desagregación por sexo y gestión del CE (estatal- no estatal).

** Estimaciones con error menor al 5%.

*** Estimaciones con error cercanos al 10%.

**** Incluyó centros educativos de secundaria de adultos.

3. Evolución del rendimiento en el tiempo

Uno de los propósitos de evaluar periódicamente el rendimiento de los estudiantes peruanos es ofrecer al sector una medida de la evolución del sistema educativo a través de los resultados obtenidos. Para ello, desde CRECER 98, la UMC ha desarrollado diseños de evaluación que buscan garantizar la comparabilidad de los resultados entre las EN.

Dichos diseños suponen definiciones metodológicas respecto a la formulación de los instrumentos y de las muestras de evaluación.

Respecto a los instrumentos, asegurar la comparabilidad de los resultados supone, por ejemplo, disponer de un banco de ítems que formen parte de una escala que pueda proporcionar ítems de anclaje entre evaluaciones, es decir, ítems que pueden volver a ser aplicados en evaluaciones posteriores. Esta es la razón por la cual gran parte de los ítems construidos por UMC no pueden ser de dominio público y por tanto se mantienen en reserva.

En tal sentido, asegurar la comparación entre las evaluaciones supone un conjunto de decisiones que debieran ser consideradas en las aplicaciones futuras:

- Mantener los dos grados terminales de cada nivel de la educación básica (sexto de primaria y quinto de secundaria)⁸
- Asegurar la evaluación del “nivel de desempeño suficiente” que define el nivel de logro esperado para el grado evaluado de acuerdo a la estructura curricular vigente.
- Mantener como áreas fijas de evaluación “matemática” y “comunicación”, en tanto son consideradas las áreas que ofrecen los códigos necesarios para otros aprendizajes. Dentro de estas áreas, mantener, como mínimo, la evaluación de las competencias o capacidades consideradas esenciales y factibles de ser medidas desde este tipo de evaluaciones. Estas son, en comunicación: Comprensión Lectora y Producción de Textos; y en matemática: Resolución de Problemas y Manejo de Algoritmos (y, del 2004 en adelante, Comunicación Matemática).
- Asegurar la comparabilidad de las muestras nacionales.

Estas consideraciones permitirán contar con los siguientes indicadores de evolución:

- Puntaje nacional promedio en cada área y competencia en la escala construida a partir del modelo de medición empleado.
- Porcentaje de estudiantes de los grados terminales de cada nivel que alcanza el nivel esperado (“nivel suficiente”) de desempeño en las principales competencias evaluadas arriba señaladas.

4. Evaluación Nacional 2004

Comparaciones en la EN 2004

Como puede deducirse del apartado anterior entre las preguntas centrales de investigación de la próxima EN están:

“¿Cuál es la evolución del rendimiento promedio de los estudiantes evaluados en la EN 2004 respecto de los evaluados en las EN 1998 y EN 2001?”

“¿En el 2004 aumentó el porcentaje de estudiantes que se ubican en el nivel suficiente respecto a la EN 2001?”

Para contestar estas preguntas será necesario respetar los criterios señalados en el apartado anterior.

Si bien la literatura internacional sugiere que el rendimiento promedio de los estudiantes no varía en corto plazo, la comparación entre 2001 y 2004 (sexto de primaria en las áreas de comunicación y matemática) y sobre todo la comparación entre 1998 y 2004 (sexto de primaria y quinto de secundaria en las áreas de comunicación y matemática) debería arrojar resultados que permitan evaluar el impacto en el aprendizaje de los estudiantes de las acciones que se han llevado a cabo desde el sistema para mejorar la calidad de la educación.

Otro tipo de comparación que se puede hacer al interior de las EN son las comparaciones de los rendimientos de grupos de estudiantes de grados próximos o cercanos. Este tipo de comparaciones ofrece información sobre los aprendizajes dentro del sistema y el posible valor agregado de más años de escolaridad. En la EN 2004 se realizará esta comparación en los grados más cercanos entre sí, a saber, tercero y quinto de

⁸ En la EN 2001 se reemplazó la evaluación del quinto de secundaria por el cuarto grado de secundaria, debido a que en ese momento se esperaba la generalización del proyecto piloto del bachillerato que establecía al cuarto grado como el último de la secundaria.

secundaria; en primaria, en cambio, la distancia entre los grados incluidos en la evaluación es tal que no es posible hacer una comparación entre los mismos.

Las áreas evaluadas

La EN 2004 contempla la evaluación de las áreas de matemática y comunicación, tradicionalmente consideradas en las evaluaciones de sistema (nacionales e internacionales) por ser las áreas básicas e instrumentales sobre las cuales se construyen otros aprendizajes. Adicionalmente, la EN 2004 considera la evaluación del eje curricular de formación ciudadana debido a que en el Perú la ausencia de una cultura consolidada de convivencia y participación democráticas se ha tornado en uno de los grandes problemas que es necesario enfrentar. El Ministerio de Educación ha asumido esta responsabilidad y la expresa a través del primer objetivo estratégico de sus lineamientos de política educativa 2001-2006: "Formar niños, niñas, jóvenes y adultos como personas y ciudadanos capaces de construir la democracia, el bienestar y el desarrollo nacional, armonizando este proyecto colectivo con su propio proyecto personal de vida"⁹, de la misma forma, el Acuerdo Nacional resalta la importancia de estos temas en su Décimo Segunda Política de Estado en la que establece "garantizar el acceso universal e irrestricto a una educación integral (...) que afiance los valores democráticos y prepare ciudadanos y ciudadanas para su incorporación activa a la vida social..."¹⁰.

Los logros de aprendizaje de los estudiantes

Las preguntas de investigación con respecto al rendimiento estudiantil que la EN 2004 se plantea responder son las siguientes:

"¿Qué nivel de desempeño muestran los estudiantes peruanos de segundo grado de primaria, de sexto de primaria, tercero de secundaria y quinto de secundaria en Comunicación y Matemática respecto a lo esperado en el currículo oficial?", y

"¿Qué nivel de desempeño muestran los estudiantes peruanos de sexto de primaria y quinto de secundaria en formación ciudadana respecto a lo esperado en el área de Personal-social / Desarrollo social del currículo oficial?"

Siguiendo con las tendencias de las EN anteriores y de otras evaluaciones internacionales, se ha incluido en la muestra a estudiantes del último grado de cada nivel, primaria y secundaria, y un grado adicional de cada nivel. Los grados adicionales elegidos esta vez son segundo de primaria y tercero de secundaria, por ser ambos fin del primer ciclo en su respectivo nivel.¹¹ Los datos de fin de nivel y de fin de ciclo deberían dar una visión de logros de ciclo, tal como se plantea en el currículo, así como del progreso de los estudiantes peruanos. Con la incorporación del segundo grado en la EN 2004, se completa la información del rendimiento de los estudiantes de los tres ciclos de la primaria.

De modo similar a la EN 2001, la propuesta de la EN 2004 busca estimar la habilidad de los estudiantes al enfrentarse a tareas de distinta complejidad. Así, la escala construida con preguntas de diferentes niveles de dificultad permitirá reportar los desempeños de los estudiantes y su nivel de adecuación al ciclo evaluado.

Los factores que influyen en los logros de aprendizaje

Con respecto a los factores asociados al rendimiento se plantea la siguiente pregunta de evaluación:

⁹ Ministerio de Educación (2001) *Lineamientos de Política Educativa 2001-2006*.

¹⁰ Acuerdo Nacional (2002).

¹¹ Cabe señalar, que si bien recientemente se ha publicado el Programa Estratégico de Desarrollo Curricular para el nivel secundario, en el que se establece que el tercer grado es el inicio del segundo ciclo de secundaria, cuando la UMC inició el diseño de la EN 2004, la propuesta curricular vigente establecía a dicho grado como fin del primer ciclo.

“¿Qué factores escolares y extraescolares están relacionados con los aprendizajes de comunicación, formación ciudadana y matemática?”

La EN 2001 se centró en los temas de currículo, formación y capacitación docente y uso de materiales educativos. En base a estos se elaboraron cuestionarios, que tuvieron como sustento teórico importante la literatura internacional sobre oportunidades de aprendizaje. En la EN 2004 se mantendrá este esquema intentando incorporar nuevas formas de medir las oportunidades de aprendizaje y en general la calidad de los procesos educativos. Por otro lado, además de recoger las variables propiamente escolares o “educativas”, para poder estimar el efecto de la escuela en el rendimiento se controlará por las características socioeconómicas y familiares de los estudiantes.

Los logros de aprendizaje de poblaciones bilingües

Responder a la diversidad cultural y lingüística del Perú a través de una Educación Bilingüe Intercultural (EBI) de calidad constituye actualmente una de las prioridades establecidas en los lineamientos de política del sector educativo.

La EN 2001 constituyó el primer esfuerzo de evaluación a muestras representativas de poblaciones vernáculo hablantes en sus lenguas maternas y en castellano. Aprovechando la experiencia adquirida en dicha evaluación, el Ministerio de Educación ha optado por volver a evaluar a los estudiantes quechua hablantes (variante del sur) y a los aimara hablantes en sus propias lenguas creando instrumentos que respondan a la lógica de sus idiomas originarios. Por otro lado, es importante también evaluar a estos grupos en castellano como segunda lengua en función de los niveles de bilingüismo de la propuesta EBI y teniendo como referencia los estándares de contenido establecidos en el currículo nacional. En ese sentido, la EN 2004 se plantea las siguientes preguntas de investigación sobre este tema:

“¿Qué nivel de desempeño en comprensión y producción de textos en su lengua materna (L1) han desarrollado los estudiantes aimara y quechua hablantes (variante del sur) de sexto de primaria que participan en el programa de EBI?”,

“¿Qué nivel de desempeño en comprensión y producción de textos en castellano como segunda lengua (L2) han desarrollado los estudiantes aimara y quechua hablantes (variante del sur) de sexto de primaria que participan en el programa de EBI?”, y

“¿Cuál es el rendimiento en comprensión y producción de textos en L1 de los estudiantes aimara y quechua hablantes (variante del sur) de sexto de primaria que participan en el programa EBI en comparación con el de los estudiantes cuya lengua materna es el castellano?”.

4. Los estándares empleados en la EN 2004

En la literatura especializada se distinguen diversos tipos de estándares educativos, entre ellos los estándares de contenido y los estándares de desempeño.¹²

Los estándares de contenido establecen las competencias, capacidades y contenidos temáticos que deberían ser aprendidos o desarrollados por los estudiantes. En nuestro sistema dichos estándares se establecen a través de los diseños o estructuras curriculares.

Los estándares de desempeño definen el nivel de logro o grados de dominio que deben mostrar los estudiantes frente a los estándares de contenido para asegurar que se han cumplido los propósitos del grado o ciclo de estudio. La tarea de definir estándares de desempeño consensuados a nivel nacional es parte de la agenda pendiente que tiene el

¹² Trahtemberg (2001)

sector. Ante la ausencia de tales estándares la UMC establece niveles de desempeño a través de consultas a jueces especializados en pedagogía del área evaluada.¹³

La EN 2004, al igual que las anteriores EN, tiene a la base el currículo peruano, de forma tal que las pruebas se diseñan sobre la matriz de competencias y capacidades que representa los contenidos básicos de los grados y ciclos de los niveles educativos evaluados.¹⁴

5. El modelo estadístico de medición de la EN 2004

El modelo estadístico aplicado a la medición educativa utilizado por la UMC es una generalización del modelo de Rasch. Este modelo se utiliza para calcular los parámetros¹⁵ asociados a los ítems y a partir de éstos elaborar escalas que representan el rendimiento de los estudiantes. Estas escalas también permiten fijar puntos de corte que definen niveles de desempeño a partir de los requerimientos curriculares. Adicionalmente, el modelo Rasch permite ubicar a los estudiantes en uno de los niveles de desempeño determinados. Para lograrlo no es necesario que los estudiantes rindan todas las tareas que forman parte de la escala, basta con que se enfrenten a un grupo de ítems de la misma para poder estimar su desempeño en el conjunto. La UMC adoptó este modelo probabilístico a partir de la EN 2001, el que ha sido utilizado en otros estudios internacionales como TIMSS¹⁶ y PISA.

6. Los niveles de desempeño de la EN 2004

Tal como se refirió en apartados anteriores, el currículo al ser un estándar de contenido organizado en tres ciclos en la primaria y dos ciclos en la secundaria, indica qué deben aprender los estudiantes peruanos. Sin embargo, en un modelo de evaluación referido a criterios es preciso comparar el rendimiento de los estudiantes con un "criterio o logro" esperado para el ciclo o grado (estándar de desempeño).¹⁷

En este contexto y con el propósito de obtener información de lo que saben hacer¹⁸ los estudiantes evaluados, es necesario establecer como criterio de la evaluación, los niveles de logro referidos a la calidad de los aprendizajes de dichos estudiantes. Para ello, como se mencionó líneas arriba, los instrumentos consideran ítems con tareas de distinto nivel de complejidad con el propósito de obtener información de la habilidad de los estudiantes a través de su desempeño, es decir, de lo que demuestran poder hacer en las pruebas. La UMC reportará resultados comparables con la EN 2001 para lo que define el siguiente nivel:

Nivel Suficiente: los estudiantes que alcanzan este nivel demuestran un dominio satisfactorio, de acuerdo con el grado evaluado, de las capacidades seleccionadas referidas a una determinada competencia.

Además, se reportarán los grupos de estudiantes que se encuentran debajo de este nivel (Básico) y los que superan dicho nivel (Avanzado).¹⁹

¹³ Ver Informe técnico de la consulta sobre puntos de corte para la EN 2001 en: www.minedu.gob.pe/mediciondelacalidad/2003/pdfs_nac/informe_puntos_corte.pdf

¹⁴ La selección y justificación de las competencias y capacidades a evaluar por área y grado se encuentra en los documentos de Especificaciones de las Pruebas.

¹⁵ En este caso los parámetros son: la habilidad del estudiante y la dificultad del ítem.

¹⁶ Estudio Internacional de Tendencias en Matemática y Ciencias (TIMSS por su nombre en inglés "*Trends in International Mathematics and Science Study*", antes conocido como "*The Third International Mathematics and Science Study*").

¹⁷ Debido a las sucesivas reformas y ajustes curriculares de nuestro sistema educativo, la UMC diseña sus evaluaciones considerando aquellas competencias y capacidades básicas y comunes a todas las estructuras curriculares.

¹⁸ Después de los resultados de la EN 2001 se ha considerado importante incrementar el número de preguntas sencillas para poder reportar de mejor manera los desempeños de los estudiantes que no alcanzan el nivel suficiente para el grado.

¹⁹ Las definiciones de estos niveles se establecerán atendiendo a criterios conceptuales y empíricos, de forma tal que proporcionen información de la mayor cantidad posible de estudiantes.

7. El ciclo evaluativo de la EN

A continuación se mencionan los pasos que la UMC realiza dentro del ciclo evaluativo de la EN. Cabe señalar que la mayoría de dichos pasos, al término de todo el proceso, se encontrarán debidamente detallados en los documentos técnicos y de difusión correspondientes.

I. Definición del diseño de evaluación

- Definición de las preguntas de investigación: incluye una primera definición de las áreas y grados a evaluar, así como de los factores asociados al rendimiento sobre los que se recogerá información.
- Definición de aspectos generales del diseño de evaluación: (muestra, diseño de medición, entre otros).

II. Construcción y validación de instrumentos

2.1 Pruebas de rendimiento:

- Análisis curricular para la selección de las competencias y capacidades a evaluar.
- Elaboración del marco de trabajo y de las especificaciones de las pruebas.
- Planteamiento del modelo de evaluación: niveles de desempeño por los que se quiere reportar, criterios de complejidad, construcción de indicadores, entre otros.
- Elaboración de los ítemes.
- Aplicación(es) piloto de los instrumentos.
- Ajuste de los instrumentos (evaluación bajo criterios estadísticos de los ítemes, de sus niveles de dificultad, de sus posibles sesgos, de su comportamiento dentro de la escala, entre otros).
- Establecimiento de los puntos de corte de las escalas.
- Diseño definitivo de las pruebas para su aplicación en campo: generalmente, los instrumentos utilizados son de lápiz y papel, y consideran tareas de distinta complejidad vinculadas a las competencias y capacidades a evaluar. Estas tareas suelen ser de selección de respuesta (ítemes de opción múltiple, de apareamiento, para responder "verdadero" o "falso") así como de producción de respuesta (ítemes de respuesta abierta corta y extendida).

2.2 Instrumentos de factores asociados al rendimiento (cuestionarios a estudiantes, docentes, directores y padres de familia):

- Elección de temas a ser abordados a partir de la revisión de literatura especializada (incluyendo resultados de las evaluaciones anteriores) y análisis de políticas del sector.
- Elaboración de la justificación de la evaluación de factores asociados y diseño de un plan preliminar de análisis.
- Aplicación(es) piloto de instrumentos preliminares.
- Ajuste de los instrumentos.
- Diseño definitivo de los cuestionarios de factores asociados

III. Aplicación en campo de los instrumentos

- Montaje de la red de aplicación (incluye coordinadores, supervisores, examinadores, entre otros)
- Diseño de manuales y guías de capacitación para los diferentes miembros de la red.
- Desarrollo de la(s) capacitación(es)

- Aplicación en campo de los instrumentos (incluye monitoreo de la calidad de los procesos)

IV. Codificación y armado de las bases

- Diseño final de criterios de codificación para las preguntas abiertas de los instrumentos.
- Capacitación a codificadores
- Digitación o lectura de fichas ópticas
- Depuración de las bases (crítica)

V. Análisis y producción de informes de resultados

VI. Difusión

- Elaboración de publicaciones
- Diseño y desarrollo de talleres y presentaciones de los resultados
- Actualización de la Página WEB

Como se puede apreciar las tareas que involucran el ciclo evaluativo son de tal envergadura que demandan un tiempo considerable, esta razón sumada a la necesidad de reportar sobre la evolución de la calidad del sistema educativo en el tiempo, ha llevado al Ministerio de Educación a tomar la decisión de establecer un periodo de 3 años entre las aplicaciones.²⁰ En tal sentido, se espera que la quinta EN se realice en el 2007.

8. Tipo de información que se obtendrá y el uso que se le puede dar

Al igual que en anteriores EN, el modelo de utilización de la información es considerado en la literatura especializada como un sistema de bajas consecuencias para los usuarios y, por tanto, la orientación es básicamente formativa. La decisión de emplear un sistema de bajas consecuencias, más que a una consideración logística, obedece a que existen indicios que advierten sobre los riesgos de ligar los resultados de evaluaciones a gran escala a sistemas de premios o castigos (“altas consecuencias”).²¹

En tal sentido, el objetivo central de la EN es proporcionar información relevante sobre el sistema educativo que contribuya a tomar decisiones de política educativa y que sirva, a la vez, de insumo para las decisiones de los diferentes actores del proceso educativo. Para conseguir este propósito son centrales los procesos de análisis y difusión de los resultados obtenidos.

Las comparaciones posibles

Además de las comparaciones en el tiempo mencionadas en el apartado número 3, los datos provenientes de la EN 2004 permitirán obtener información sobre diferentes estratos o grupos poblacionales y comparar sus desempeños. Esto permitirá obtener indicadores de equidad dentro del sistema e identificar grupos en mayor riesgo de no alcanzar los objetivos educativos propuestos desde el sector.

El diseño muestral permitirá comparar los desempeños de los estudiantes de ambos sexos, de centros educativos de gestión estatal y no estatal, y de los centros educativos de ámbitos rurales y urbanos. Además, proporcionará una aproximación de los desempeños de los estudiantes a nivel departamental.

²⁰ Esto responde además a la dinámica de los cambios en los aprendizajes de los estudiantes del sistema educativo que, como se sabe, no deberían variar significativamente en el corto plazo.

²¹ Entre los riesgos principales se encuentran: el favorecer una práctica pedagógica orientada principalmente a entrenar a los estudiantes en rendir pruebas estandarizadas, el restar valor a todas aquellas dimensiones y aprendizajes que no pueden ser evaluados en este tipo de operativos, o el favorecer un movimiento selectivo de la demanda que provoca la concentración de “buenos alumnos” en “buenas escuelas” y de “malos alumnos” en “malas escuelas”, entre otros. De ahí que las evaluaciones de altas consecuencias deben ser formuladas con extrema cautela.

En primaria, adicionalmente, se reportará por característica del centro educativo (escuelas polidocente-completas vs escuelas multigrado²²).

Principales usuarios

A continuación se presenta una relación de los principales usuarios de la información que las EN provee:

- Alta dirección del Ministerio de Educación, Directores Nacionales y Regionales.
- Equipos técnicos de la Sede Central y de los Órganos Descentralizados del Ministerio de Educación: en este grupo se encuentran los responsables de la elaboración del currículo, de los materiales educativos, de los diseños de capacitación, de la planificación y gestión del sector, entre otros.
- Instituciones de la sociedad civil que intervienen en el quehacer educativo tales como redes, foros, u organizaciones no gubernamentales.
- Docentes y directivos de los centros educativos del país.
- Centros de formación docente como Institutos Superiores Pedagógicos y Universidades.
- Comunidad académica.

Para atender las demandas y necesidades de información de cada uno de los actores señalados, la UMC pondrá en marcha un plan de difusión que contemplará la producción de diferentes tipos de reportes, diseño de talleres, presentaciones, entre otros.

Por ejemplo, pensando principalmente en los usuarios de los centros educativos, institutos de formación docente y en los equipos técnicos – pedagógicos se elaborarán reportes descriptivos detallados de las tareas que los estudiantes son capaces de realizar y de los niveles de desempeño sobre los que la EN 2004 permitirá reportar. Así, el lenguaje y los énfasis temáticos se graduarán en función de las características de los públicos objetivos.

Adicionalmente, el Ministerio reconoce su obligación y compromiso de mantener informada a la opinión pública sobre la calidad del servicio que ofrece en aras de fortalecer una cultura de evaluación y calidad dentro del sector. La prensa adecuadamente informada es un elemento clave para conseguir este propósito.

²² Las escuelas multigrado incluyen a las polidocente multigrado y a las unidocentes.

REFERENCIAS BIBLIOGRÁFICAS

Secretaría Técnica del Acuerdo Nacional (2002)

Texto del Acuerdo Nacional.

Secretaría Técnica del Acuerdo Nacional: Lima.

Benavides, M. , G. Espinosa y A. Montané (2002)

Evaluación de sistema y evaluación de aula.

MINISTERIO DE EDUCACIÓN – UMC: Lima.

http://www.minedu.gob.pe/mediciondelocalidad/2003/zips/ev_sistema_ev_aula_preliminar.pdf

Cooper S. y M. Lyn (2000)

National assessment of educational progress achievement levels.

NATIONAL ASSESSMENT GOVERNING BOARD: Washington, DC

De Moura C. y M. Carnoy (1997)

Cómo anda la reforma da educação na America Latina?

FUNDACIÓN GETULIO VARGAS: Río de Janeiro.

Ministerio de Educación (2001)

Lineamientos de Política Educativa 2001-2006.

MED: Lima.

Ministerio de Educación (2002)

Fundamentación general de la Evaluación Nacional del Rendimiento Escolar 2004.

MED-UMC: Lima.

Ministerio de Educación (2004)

Programa Nacional de Emergencia Educativa 2004: Prepublicación para recoger opiniones y recomendaciones de la comunidad educativa nacional.

MED: Lima.

Trahtemberg, L. (2001)

La polémica tarea de evaluar la calidad. En: Revista Crecer N°2. MED-UMC: Lima.