



Organización
de Estados
Iberoamericanos

Para la Educación,
la Ciencia
y la Cultura

Revista Iberoamericana de Educación

Número 10 - Enero Abril 1996

Evaluación de la Calidad de la Educación

Datos Artículo

Título: Laboratorio Latinoamericano de Evaluación de
Calidad de la Educación

Autor: Juan Casassús, Violeta Arancibia, Enrique
Froemel

Laboratorio Latinoamericano de Evaluación de Calidad de la Educación

Juan Casassus
Violeta Arancibia
Juan Enrique Froemel (*)

El Laboratorio Latinoamericano de Evaluación de Calidad de la Educación es un recurso técnico a disposición de los países latinoamericanos. Constituye también un ámbito de discusión técnico-política para la problemática del aprendizaje y sus variables relacionadas.

Sus objetivos consisten en la identificación de estándares de aprendizaje escolar para la región y la apreciación del grado y nivel de logro de tales estándares en los países; el fomento del cambio educativo que permita alcanzar tales estándares; y la formación de recursos humanos que hagan posible dicho cambio.

231

1. Introducción

Para los efectos anteriores el Laboratorio deberá desarrollar las líneas de acción que se señalan a continuación:

- a. Explorar la determinación de estándares de calidad y de nivel de la Educación en la región.
- b. Determinar (medir y evaluar) y hacer seguimiento del estado de los niveles de aprendizaje escolar. También, obtener información respec-

(*) Juan Casassus es doctor en Economía de la Educación y especialista regional en Políticas, Planificación y Gestión Educativas de la Oficina Regional de Educación para América Latina y el Caribe, OREALC/Unesco.

Violeta Arancibia es PHD en Educación y docente de la Escuela de Psicología de la Pontificia Universidad Católica de Santiago de Chile.

Juan Enrique Froemel es doctor en Educación y consultor de OREALC/Unesco.

to de su calidad y del estado de variables incidentes en los países de América Latina. Lo anterior para posibilitar la valoración de la calidad y nivel de la Educación.

c. Desarrollar, establecer, mantener y actualizar bases de información respecto de la situación del aprendizaje escolar y variables incidentes en los países.

d. Experimentar con enfoques innovadores que permitan el cambio de los métodos pedagógicos, en pro del aumento y mejoramiento del aprendizaje escolar en la región.

e. Difundir los hallazgos respecto del aprendizaje escolar y de sus variables incidentes en los países participantes. Asimismo, disseminar los resultados de la experimentación en los mismos países.

f. Realizar actividades de formación y capacitación del personal de los sistemas de medición y evaluación del aprendizaje escolar en los países participantes.

2. Antecedentes

En el marco de la Recomendación para la ejecución del Proyecto Principal de Educación en América Latina y el Caribe, en el período 1993-1996, se establece que tanto en el contexto externo vinculado a la Educación como al interior de los sistemas educativos, se ha generado un conjunto de condiciones, posibilidades y necesidades. Estas, a su vez, generan nuevas demandas para la superación del endémico desfase entre las características del sistema educacional y los requerimientos individuales y sociales.

En el documento aludido se analizan las posibles causas de que las reformas educativas en América Latina no hayan alcanzado los resultados esperados. Se señala, entre otras, la inestabilidad de las políticas educativas en el tiempo, debido a su dependencia de autoridades gubernamentales sucesivas. También se indican limitaciones con respecto del volumen y pertinencia de la información disponible para diseñar estas reformas. Finalmente, se apunta hacia la dificultad de que las medidas adoptadas produzcan efectos en el aula.

Frente a la situación descrita se sugieren algunos caminos de solución. En relación con la primera causa, se plantea la necesidad de establecer convergencias y acuerdos básicos en la formulación de políticas. En cuanto al problema de la información, se sugiere la identificación de sistemas adecuados de recolección y procesamiento de la

misma. Finalmente, en cuanto a la real aplicación de las medidas adoptadas a nivel de aula, se recomienda que el diseño de la reforma incorpore una componente de conducción y seguimiento sistemático de la aplicación y efecto de las medidas.

En tal contexto, es posible afirmar que el Laboratorio de Evaluación de Calidad de la Educación, constituyendo un proyecto de cooperación regional, se enmarca en algunos de los objetivos del Proyecto Principal de Educación en América Latina y el Caribe. Ellos son:

Primer objetivo: la escolarización de todos los niños en edad escolar y la duración de la Educación general mínima.

Segundo objetivo: la acción de alfabetización y los servicios educativos para los adultos.

Tercer objetivo: la calidad y eficiencia de los sistemas educativos.

En el marco de los objetivos planteados, el Laboratorio constituye además una herramienta útil para promover una gestión educativa responsable de sus resultados y logros, un tema que deberá cobrar creciente relevancia en nuestros países en el futuro cercano. Este aspecto constituirá un factor crucial en la modernización de la administración educativa.

Dentro de ese nuevo estilo de gestión, es preciso incorporar como partes integrantes de los planes y programas los niveles de logro y los indicadores de la calidad del aprendizaje escolar que se esperan alcanzar en cada caso. Consecuentemente, es preciso aplicar en forma regular procedimientos de determinación (medición y evaluación) de los resultados obtenidos en cada programa. Estos últimos, además, deben retroalimentar y perfeccionar permanentemente el sistema de gestión.

Los principales ejes de acción para el mejoramiento de los niveles y la calidad del aprendizaje han de ser el institucional y el pedagógico.

En cuanto al eje institucional, interesa destacar el rol fundamental que se asigna a los Ministerios de Educación en la incorporación de estándares nacionales y sistemas de medición y evaluación de productos del proceso educativo. Se insta a los Ministerios a introducir estándares cada vez más exigentes para cada grado, dirigidos al desarrollo de aprendizajes de nivel superior, tales como el pensamiento lógico, como así también las competencias y habilidades para la resolución de problemas. En este último aspecto, el Laboratorio deberá constituir una fuente válida para la exploración de estándares relevantes a disposición de los países involucrados.

Un sistema educativo que pretenda anticiparse a las demandas futuras requiere, además, desarrollar una política que fomente las innovaciones. Consecuentemente, los Ministerios deben incentivar, conducir, evaluar, sistematizar y divulgar las innovaciones educativas. Para esto último, el Laboratorio surge como un elemento útil y propicio.

Con respecto al eje pedagógico, se señala que el aprendizaje de la lectura, la escritura y la matemática son los indicadores fundamentales de la calidad y del nivel de la Educación para el ciclo formado por los tres o cuatro primeros grados de enseñanza. Por su carácter instrumental y formativo, estos aprendizajes facilitan el acceso a la cultura y al desarrollo personal y son la base para un aprendizaje continuo.

3. Objetivos

3.1. Objetivos generales

A. Identificará estándares de aprendizaje escolar para la región y la apreciación del grado y nivel de su logro.

El Laboratorio identificará estándares de aprendizaje de acuerdo a las evidencias internacionales disponibles. También diseñará, elaborará y pondrá en marcha un sistema de apreciación (medición y evaluación) de los niveles y calidad del aprendizaje en lenguaje y matemática, para su contrastación con los estándares.

Para juzgar la calidad y el nivel de la Educación se requiere de una multiplicidad de índices. Algunos de ellos pueden desarrollarse en forma relativamente simple a partir de datos censales y de algunas de las estadísticas básicas publicadas periódicamente por los sistemas educacionales. Otros, en cambio, son difíciles y costosos de recoger y formarán parte de un programa permanente de recopilación.

El Laboratorio desarrollará y operará un sistema que permita hacer seguimiento de los índices relacionados con el aprendizaje en los sistemas escolares involucrados en el proyecto.

Finalmente, el Laboratorio establecerá un banco de información que permita conservar, actualizar y tener disponible aquella respecto del aprendizaje y variables incidentes, recopilada en los países

B. Fomentará el cambio educativo que permita alcanzar los estándares.

El Laboratorio estará orientado no sólo a producir cambios en la Educación, sino también hacia la comprensión y explicación de las

variables que inciden en dichos cambios, de sus efectos en los alumnos y de sus consecuencias a largo plazo para la sociedad. Un instrumento crucial para el cambio efectivo surge de la experimentación de nuevas soluciones a los problemas del aula.

El Laboratorio diseñará y desarrollará investigaciones en las variables más influyentes sobre la calidad y el nivel de la Educación y que sean más relevantes para el diseño de políticas educacionales.

Al servicio del primer objetivo y también de éste, el segundo, surge la necesidad de establecer un sistema de entrega y uso de la información para las distintas audiencias de los sistemas escolares de la región. Este sistema incluirá la difusión de los productos de la experimentación, como fuente para la innovación.

C. Formará recursos humanos que permitan lograr el cambio.

Sin perder de vista que el Laboratorio será el elemento pivotal para la determinación del nivel y calidad del aprendizaje en los países involucrados, se requiere de competencia técnica local en esos mismos países.

Lo anterior por dos razones. La primera dice relación con la necesidad de contar con información objetiva, válida y confiable para cumplir las funciones del Laboratorio. La segunda razón se relaciona con la necesidad de que los países puedan perfeccionarse en la satisfacción de sus propias necesidades de determinación (medición y evaluación) de los niveles y calidad del aprendizaje.

3.2. Objetivos específicos

A. Con respecto al primer objetivo general:

a. Propiciará un acuerdo entre los planificadores y los gestores de políticas educacionales de los países en cuanto a los estándares mínimos para el nivel y calidad del aprendizaje escolar.

b. Identificará un conjunto de operacionalizaciones que den cuenta del logro de los estándares .

c. Definirá el conjunto de variables internas y externas al sistema escolar, que afectan o explican las diferencias en el logro de contenidos y destrezas.

d. Construirá instrumentos para la apreciación (medición y evaluación) de los niveles y calidad del aprendizaje en lenguaje y matemática.

e. Construirá instrumentos para recoger información acerca de las variables internas y externas al sistema escolar, antes mencionadas .

f. Validará y aplicará experimentalmente los instrumentos, a fin de determinar su posibilidad de generalización a los países de América Latina y recopilará experiencias de su administración.

g. Aplicará los instrumentos validados a muestras representativas de cada país participante en el Laboratorio.

h. Procesará la información proveniente de los distintos países y la contrastará con los estándares.

i. Diseñará un plan de difusión de los resultados que permita sensibilizar gradualmente a los sistemas educacionales de los diferentes países, a los actores del proceso educativo y a la opinión pública en general, respecto de los sistemas de apreciación (medición y evaluación) y de su impacto en el mejoramiento del nivel y la calidad de la Educación.

j. Difundirá los resultados obtenidos para colaborar en la orientación de las decisiones de políticas educacionales.

k. Estructurará y mantendrá un banco de información con los instrumentos aplicados y con los resultados obtenidos en cuanto a niveles y calidad del aprendizaje escolar y variables incidentes en éste.

B. Con respecto al segundo objetivo general:

a. Desarrollará experimentación que contribuya a orientar a los países respecto de cómo alcanzar gradualmente los estándares fijados.

b. Elaborará un manual de uso de la información obtenida respecto del aprendizaje y variables incidentes, que estimule la generación de innovaciones educativas en la región.

c. Recogerá y difundirá las experiencias innovadoras que se produzcan en los países participantes, a partir del programa aplicado por el Laboratorio.

d. Editará una publicación anual que estimule el conocimiento y la transferencia de experiencias innovadoras de la región.

C. Con respecto al tercer objetivo general:

a. Desarrollará cursos, talleres y seminarios de capacitación y perfeccionamiento en evaluación y medición educacionales, en la difusión, almacenamiento y uso de la información generada a través del Laboratorio.

b. Posibilitará estadías en el Laboratorio para profesionales de los sistemas de medición y evaluación de los países latinoamericanos.

c. Ofrecerá y eventualmente entregará asistencia técnica en su temática a los países que la soliciten.

4. Organización

Laboratorio Latinoamericano de Evaluación de Calidad de la Educación.

Participantes nacionales. Coordinadores de los Ministerios de Educación de: Argentina, Bolivia, Brasil, Colombia, Chile, Costa Rica, Cuba, El Salvador, Honduras, México, Paraguay, Perú, República Dominicana y Venezuela.

Consultores Regionales.

Secretaría Ejecutiva.
UNESCO/OREALC.

Consultores Internacionales.

5. Justificación de un programa de medición y evaluación

La apreciación del nivel y la calidad de la Educación por medio de la medición y evaluación, cobran cada vez mayor importancia, en la medida en que los países reconocen las ventajas de establecer y comparar sus rendimientos dentro de un contexto internacional.

Las estadísticas que los países usualmente recolectan no son exhaustivas en materias de Educación. La mayoría de los esfuerzos se dedica principalmente a describir y cuantificar variables de entrada. No se ha otorgado virtualmente ninguna atención a documentar cómo las escuelas funcionan o qué y cuánto aprenden los estudiantes. Incluso, la mayoría de los datos que pueden ayudar a definir el uso de los recursos o a establecer la efectividad de los recursos para la Educación, no están a la mano.

En lo metodológico, se avizora que en la presente década se producirá una creciente y gradual complementación entre los enfoques cuantitativos y los cualitativos, para la determinación de los logros del aprendizaje, conforme las autoridades y planificadores centren su atención en nuevos medios de mejorar la enseñanza.

Además de la observación del desarrollo de los conocimientos y destrezas requeridos para que los jóvenes sean parte activa de la cultura contemporánea y de posibilitar su inserción en el mercado de trabajo, otra razón para realizar estudios sobre el rendimiento escolar es la búsqueda de una pedagogía más eficaz. Los estudios internacionales facilitan el estudio comparado de la eficiencia educativa en varios sistemas simultáneamente. Además, permiten generar informes nacionales respecto del

aprendizaje, centrados en eventuales diferencias basadas en el género, la condición urbana o rural, entre escuelas polivalentes o tradicionales y otras. Tales estudios aportan abundante información sobre la Educación en un país, su rendimiento y las variables incidentes en éste.

Por lo general, los estudios internacionales comparados incluyen a millares de estudiantes que se someten a pruebas estandarizadas, previamente validadas, basadas en sus programas o adaptadas a otras necesidades intra o extraescolares. Las pruebas tienen por objeto determinar el rendimiento del alumno en forma tal que pueda cuantificarse, adaptarse y compararse con el de otros países y correlacionarse con otros indicadores de rendimiento, tales como la estructura de los planes de estudio, la idoneidad del profesorado, las tecnologías de enseñanza, los recursos escolares, etc. La finalidad de estos estudios es ofrecer un modelo lo suficientemente amplio como para que puedan apreciarse los efectos globales de la enseñanza sobre el aprendizaje.

6. La calidad de la educación: una discusión conceptual

La primera consideración antes de entrar de lleno en esta discusión afecta precisamente a la denominación del Laboratorio. Al respecto, aunque se sustenta que existen, al menos, dos ámbitos bajo los cuales evaluar la Educación, uno cualitativo y por ende ligado a la calidad y otro cuantitativo y relacionado con el nivel de logros en la Educación, se ha escogido mencionar en su denominación solamente la dimensión cualitativa. Lo anterior, por cuanto existe un uso generalizado de referirse a esta dimensión aun cuando se haga referencia a ambos, la calidad y el nivel.

En general se suele abordar el concepto de calidad a través de dos aproximaciones diferentes: una de ellas procura discutirlo y definirlo en forma constitutiva o conceptual y la segunda se centra en la operacionalización de la calidad y se refiere más propiamente al nivel de logros en Educación.

Este último enfoque emplea, a menudo, el término calidad como sinónimo de otros conceptos afines, tales como efectividad y eficiencia. Bajo esta perspectiva se alcanzan opciones para la determinación de la calidad de la Educación y la formulación de políticas para su mejoramiento.

Los aspectos más relevantes y consensuales de la calidad de la Educación son los siguientes:

a. La necesidad de efectuar apreciaciones diagnósticas acerca de los niveles y calidad existentes, lo que implica la determinación del nivel de logro de los objetivos educacionales y la estimación de las habilidades y destrezas adquiridas por los educandos en la escuela.

b. Aunque la calidad no es posible determinarla con exactitud, principalmente por las limitaciones de las formas de operacionarla utilizadas en las diferentes aproximaciones al problema, las informaciones disponibles acerca de la repitencia, deserción, relevancia curricular y magnitud del efecto de la Educación en el desarrollo social, apuntan a que la calidad de los sistemas educativos latinoamericanos es deficiente.

c. La calidad aparece ligada al problema de equidad. En efecto, la calidad de la Educación que reciben los sectores más pobres es notoriamente más baja, acentuando las diferencias sociales y económicas. Las apreciaciones de la calidad pueden ayudar a superar el problema de la falta de equidad en el servicio educativo, al identificar las escuelas más carentes y que consecuentemente requieren acciones concretas que les permitan mejorar la calidad de la Educación entregada a los sectores más desfavorecidos en lo cultural, social y económico.

Es preciso tener en cuenta, eso sí, que antes de la aplicación de medidas que, por la vía de asegurar una determinada calidad de Educación busquen aproximarse a la equidad, existen otros pasos imprescindibles.

Entre esos elementos se halla la necesidad de una gestión de los sistemas educativos que eleve la eficiencia en el manejo de los recursos. Se trata de procurar una mejor canalización del gasto en Educación que minimice aquella porción que se pierde y que va, por ende, en desmedro de la equidad.

Las autoridades encargadas de las decisiones educacionales deben ser sensibles a la idea de que son responsables («accountable») de favorecer o desfavorecer la equidad, por la vía de su mayor o menor eficiencia. Los resultados de la determinación de los niveles y de la calidad del aprendizaje constituyen también un elemento válido para percibir la eficiencia de la gestión de los sistemas y su efecto sobre la equidad.

La orientación conceptual del Laboratorio es la comprensión del significado de la calidad en estrecha relación con el nivel de logro de los objetivos educacionales, en el marco de los programas oficiales de estudio, tomando en cuenta las variables de insumo y especialmente las de proceso. Lo anterior, de modo de poder explicar las diferencias detectadas a diferentes niveles de agregación de los resultados.

Por tanto, interesa determinar los niveles del aprendizaje escolar, en cuanto contenidos, procesos cognoscitivos y desarrollos afectivos, así como la capacidad explicativa de las variables contextuales que inciden en las diferencias de logros de aprendizaje alcanzados.

7. Esquema interpretativo: variables

El concepto subyacente al Laboratorio implica que la determinación del nivel y de la calidad de la Educación consisten básicamente en la definición de un conjunto de variables que proporcione, en forma sistemática, un cuadro confiable y válido acerca del estado de los sistemas educativos y que pueda ser utilizado para colaborar en la orientación de las políticas y acciones de mejoramiento. Que al mismo tiempo permita estimar válidamente el mejoramiento que experimentan los sistemas como consecuencia de dichas acciones. Por lo tanto, el proceso de determinación de niveles y de calidad aborda tres aspectos, a saber: la selección de las variables, la organización de ellas en un sistema y la información de los resultados de su valoración. La operacionalización de los conceptos de nivel y de calidad de la Educación a través de variables, permite analizar el sistema educacional e informar acerca del funcionamiento del mismo, proporcionando claves para la fundamentación de la toma de decisiones dirigida al mejoramiento del sistema.

El tipo de variables que se ha seleccionado implica que éstas cumplan al menos con una de las siguientes características:

240 a. Describir el funcionamiento del sistema en función del logro de metas educacionales.

b. Describir los elementos del sistema que explican las desigualdades en los resultados y el impacto del sistema educacional, a diferentes niveles de agregación. Este tipo de variable puede aportar a la predicción del funcionamiento futuro del sistema.

c. Proporcionar información acerca de los rasgos más permanentes del sistema. Este tipo de información puede ayudar a comprender mejor cómo funciona el sistema y a estimar el efecto de los cambios en el tiempo.

Las variables aisladas proporcionan información muy limitada y, por lo tanto, es necesario organizarlas en un esquema que proporcione la representación más válida y coherente posible de la condición en que se encuentra un sistema o subsistema. Una aproximación tal suele proporcionar mejor información que el uso de la simple agregación de las variables.

La selección de las variables requiere el desarrollo de un marco conceptual basado en las necesidades planteadas por las autoridades educacionales y los educadores de cada país y en la investigación educacional.

La intención es arribar a un esquema que se base en la integración ponderada de un grupo de variables. Una determinación «a priori» de las ponderaciones de las variables considera los tres criterios siguientes:

- La importancia y utilidad, de general aceptación, para cada variable.
- La calidad técnica posible para la obtención de sus valores.
- El menor costo relativo de su apreciación.

Adicionalmente, se ha considerado reducir el número de indicadores en pro de la simplicidad del esquema.

A su vez, la organización de un sistema de variables implica decisiones acerca de los métodos y formas de recolección de la información, a la vez que la definición de patrones para producir datos interpretables y comparables.

El sistema de variables representativo del nivel y de la calidad de la Educación que se ha escogido, se ha construido tomando en cuenta los estudios internacionales en los cuales tales variables han tenido un potencial explicativo sobre los resultados del proceso educativo, los objetivos del laboratorio y la factibilidad de ser estimadas en forma masiva.

El esquema interpretativo propuesto relaciona seis áreas de variables incidentes en el aprendizaje escolar con dos tipos de incidencia (de insumos y de proceso). También considera los niveles y la calidad del aprendizaje escolar como un conjunto de variables de producto del sistema de Educación formal, al cual convergen las anteriores.

Es necesario señalar, eso sí, que todo esquema interpretativo, en cuanto paradigma, es una sobre-simplificación del esquema conceptual subyacente. Por lo anterior es imprescindible advertir que entre cada par de niveles, insumos-procesos, procesos-productos, productos-insumos, existen necesariamente retroalimentaciones. En el esquema, tales componentes se obviaron en beneficio de la claridad y simpleza del esquema.

Las seis áreas de variables incidentes son: la política social-educativa, la familia del estudiante, la escuela, el alumno, el currículum y el profesor.

Las variables se han organizado de acuerdo al esquema interpretativo propuesto, en las siguientes categorías generales:

A nivel de insumos:

a. Política social-educativa:

- Gasto en Educación del país
- Gasto en Educación por alumno
- Escolaridad obligatoria

b. La familia del estudiante:

- Composición del grupo familiar
- Nivel ocupacional de los padres
- Nivel educacional de los padres
- Equipamiento del hogar
- Bienes y recursos culturales

c. La escuela:

- Ubicación geográfica (rural, urbana)
- Dependencia administrativa (particular, municipal, otra)
- Calendario escolar
- Razón profesor por cantidad de alumnos
- Razón personal directivo por cantidad de alumnos
- Razón personal administrativo por número de alumnos
- Número de estudiantes por curso
- Número de alumnos por escuela
- Infraestructura básica (bancos, salas de clase, gimnasio, biblioteca)
- Recursos instruccionales (textos para el profesor, textos para los alumnos, medios y materiales de apoyo a la enseñanza, audiovisuales, etc.)
- Sistema de incentivos para directivos y profesores

d. El alumno:

- Género
- Edad
- Etnia
- Nivel socioeconómico
- Historia escolar (rendimiento, variables afectivas)

e. El currículum planificado:

- Ciclos
- Asignaturas
- Contenidos
- Objetivos

f. El profesor:

- Género
- Edad
- Etnia
- Formación previa
- Formación de postgrado
- Perfeccionamiento en servicio
- Años de experiencia
- Remuneración
- Jornada de contratación
- Compromiso y satisfacción personal

A nivel de procesos

a. Política social-educativa: Gestión institucional

- Sistema de administración (centralizado, descentralizado)
- Sistema de asignación de recursos
- Incentivos a la innovación en las escuelas
- Supervisión y apoyo técnico a las escuelas

b. La familia del estudiante: Compromiso educacional familiar

- Hábitos, rutinas y participación de los niños en las tareas hogareñas
- Guía, apoyo y estímulo académicos entregados a los hijos
- Expectativas y aspiraciones de los padres respecto a la Educación de los hijos
- Participación en la escuela
- Satisfacción con el establecimiento al que asisten sus hijos

c. La escuela: Gestión escolar

- Clima organizacional
- Liderazgo académico del Director
- Rotación de personal docente
- Ausentismo laboral
- Ejercicio de funciones y atribuciones del Director
- Expectativas del Director
- Atribuciones del Director
- Niveles de participación de la comunidad educativa

d. El alumno: Hábitos y destrezas

- Asistencia a clases
- Puntualidad
- Tiempo en la tarea
- Participación en clases

e. El currículum realizado:

- Número de días efectivos de clases
- Número de horas efectivas de clases
- Cumplimiento de metas
- Impedimentos o dificultades en su desarrollo

f. El profesor: Gestión pedagógica

- Estilo de enseñanza
- Empleo de medios y materiales de apoyo a la enseñanza
- Asignación de tareas escolares
- Revisión de las tareas asignadas
- Oportunidad de aprendizaje
- Distribución del tiempo de los profesores
- Expectativas de los profesores
- Atribuciones de los profesores
- Frecuencia de las evaluaciones
- Disciplina escolar

A nivel de productos: Indicadores de nivel y calidad del aprendizaje escolar

- Logro de contenidos mínimos en matemáticas
- Logro de contenidos mínimos en lenguaje
- Nivel de destrezas cognoscitivas de cálculo
- Nivel de destrezas cognitivas lingüísticas
- Nivel de autoestima académica
- Años para completar el primer ciclo básico
- Tasa de deserción
- Tasa de repitencia

El Laboratorio pretende llegar a estimar estas variables, tratando sobre todo de establecer empíricamente las interrelaciones explicativas que permitan validar las hipótesis implícitas en el esquema interpretativo propuesto.

Las hipótesis básicas que se pretende validar se expresan en términos generales e integradores de varias sub-hipótesis posibles, relacionando las áreas de variables incidentes del esquema interpretativo propuesto y las variables de producto de nivel y de calidad del aprendizaje señalados.

Estas hipótesis generales e integradoras son las siguientes:

Hipótesis 1.

La calidad de los aprendizajes de los niños en la escuela depende especialmente de la cantidad de insumos que la política social-educativa asigna al sistema educativo y del apoyo técnico que la gestión institucional proporciona a las escuelas para llevar a cabo proyectos innovativos.

Hipótesis 2.

El aprendizaje escolar depende parcialmente de las características culturales, sociales y económicas de la familia de los estudiantes y de la calidad de los procesos ambientales de su hogar, como asimismo de su interacción, los que representan el grado de compromiso de la familia en relación con las actividades escolares de sus niños.

Hipótesis 3.

El aprendizaje escolar está asociado a las características de la escuela y a los recursos con que cuenta, así como a la calidad de la

gestión liderada por el personal directivo y al grado de participación de la comunidad escolar en ella.

Hipótesis 4.

Los niveles y la calidad del aprendizaje escolar dependen parcialmente de las características culturales, sociales, económicas y de ingreso a la escuela de los niños, así como de su historia escolar y de sus hábitos y destrezas de estudio.

Hipótesis 5.

El aprendizaje escolar depende parcialmente de las características del currículum planificado y del currículum realizado, particularmente en lo que se refiere al cumplimiento de las metas establecidas y a la continuidad de su desarrollo.

Hipótesis 6.

Los niveles y la calidad del aprendizaje de los niños en la escuela dependen parcialmente de las características profesionales del profesor y de la calidad de su gestión pedagógica en la conducción del proceso de enseñanza-aprendizaje.

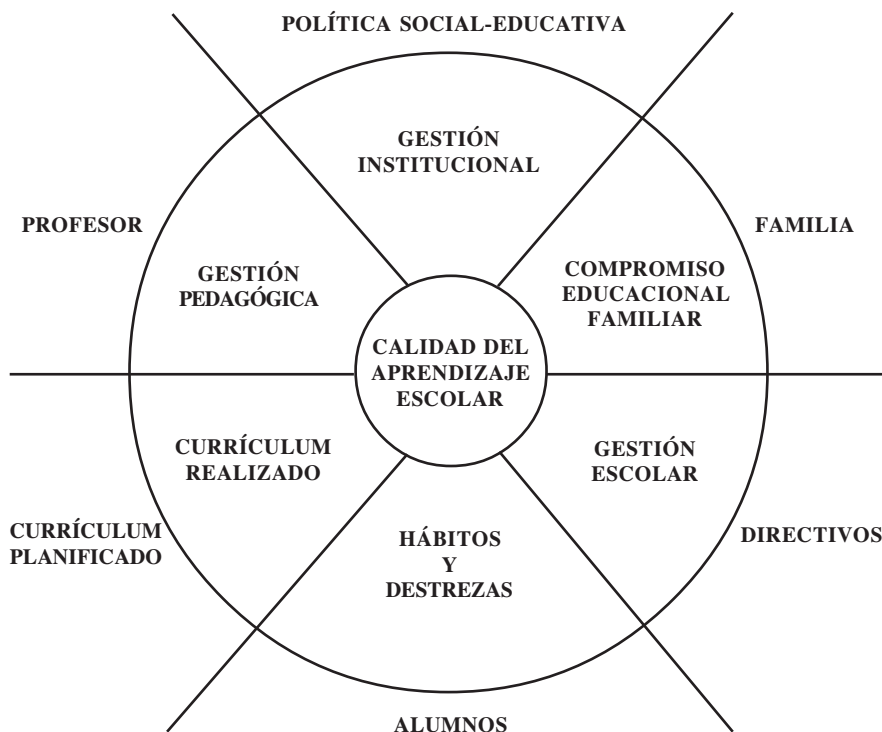
Una hipótesis todavía más general subyace al esquema propuesto. Esta plantea que el grado de incidencia de las variables de insumo en el aprendizaje escolar es menor que el de las variables de proceso, en cada una de las seis áreas indicadas.

Lo anterior requiere de mediciones objetivas, válidas y confiables, que permitan explorar las hipótesis propuestas, usando los procedimientos estadísticos multivariados adecuados.

Los procedimientos estadísticos apropiados deberán incluir, además de los descriptivos, el análisis de ítems en cada uno de los instrumentos empleados, análisis de varianza y covarianza, regresión múltiple (eventualmente con partición de varianza en bloques de variables), y análisis de trayectorias (path analysis).

Para estos procedimientos estadísticos será necesario diferenciar las variables individuales (tales como el alumno, el profesor, el director) de las variables colectivas (características de la escuela, por ejemplo). Del mismo modo, deben identificarse los niveles de agregación y de análisis (medidas de los alumnos y promedio del curso, por ejemplo).

8. Esquema interpretativo: representación gráfica



9. Sobre los estándares

Origen del tema

La preocupación por la calidad en Educación a inicios de los noventa generó conversaciones y discusiones acerca de nuevos temas. Estas tocaban aspectos tales como la atención a los resultados y el logro, énfasis en la gestión, focalización en los procesos, concertación con los usuarios.

Uno de esos temas se refirió al establecimiento de una referencia que pudiera ser utilizada en la formulación de los juicios acerca de la calidad. Esa referencia es lo que aquí denominamos un estándar. En PROMEDLAC V, los ministros de Educación recomendaron que en cada país se determinaran estándares nacionales. También recomendaron como deseable la determinación de estándares regionales.

Estándares regionales, ¿para qué?

Los estándares son criterios útiles en la formulación de juicios acerca del estado de una determinada área de la Educación.

Ellos sirven para:

- compararse y situar a un país en relación con los otros países,
- fomentar la regionalidad facilitando la comunicación, la movilidad de personas entre países y acentuar la cultura latinoamericana,
- facilitar el desarrollo de la acreditación académica entre los países,
- identificar las áreas que requieren cambios para el logro de los estándares, mejorar los niveles académicos y abrir posibilidades de renovación pedagógica,
- identificar lo que los niños en tercer y cuarto grado (en el caso presente) pueden aprender y hacer en un área determinada de la educación,
- determinar una referencia en el tiempo.

248

Estándares de contenido y de competencias

Si bien en distintas esferas de la vida se acostumbra a hacer referencia a estándares: estándares en edificios, estándares en cuanto a los automóviles, salud, hoteles, comida, más aún estándares en cuanto a la seguridad de los automóviles, a la comodidad de los hoteles o a la salubridad y gusto de los alimentos, etc. Sin embargo, la discusión acerca de los estándares en Educación es reciente y no existe aún idea consensuada acerca de lo que ellos pueden ser. En la práctica coexisten interpretaciones diversas acerca de lo que se dice cuando se hace referencia a los estándares en Educación. Por ello, existen diversas visiones acerca de lo que deberían ser los estándares, sean ellos nacionales o regionales.

Un primer paso es estar de acuerdo en que los estándares se refieren a un cierto nivel y calidad de la Educación. Mas allá de ello, existen visiones ligadas a dos temas principales: estándares basados en los objetivos pedagógicos y aquellos vinculados a la medición/evaluación.

Las visiones que vinculan el estándar con los objetivos educacionales tienen que ver con el dominio sustantivo y se los denomina estándares de contenido. Entre ellas están básicamente:

i) aquellas que se refieren a un conjunto de conocimientos o habilidades fundamentales deseables para una determinada área, y,

ii) aquellas que constituyen una formulación precisa acerca de lo que deben saber y hacer los niños de determinado nivel (edad o grado o terminal) en relación con un área específica.

Las que vinculan el estándar con la medición son aquellas denominadas de competencias, es decir, aquellas que entregan información acerca de lo que puede ser considerado inadecuado, aceptable, competente o sobresaliente. Entre ellas están:

iii) aquellas que se refieren al nivel de logro que deben alcanzar los alumnos. Este nivel (normalmente expresado en números) o estándar puede ser considerado como el saber mínimo (lo básico) que debe lograr un alumno.

iv) aquellas en las cuales el estándar es un conjunto de puntos ubicados en una escala de logro indicando distintos niveles de competencia. Estándares de nivel múltiple.

El primer tipo de estándar busca formular aseveraciones relativas a un área del conocimiento o del hacer. Dichas aseveraciones se pueden referir a conceptos claves, a contenidos específicos o a destrezas.

Sin embargo, no es fácil distinguir entre las aseveraciones relativas a los estándares ya aquellas relativas a los objetivos pedagógicos. Ellas se diferencian entre sí por el nivel genérico de las aseveraciones y también porque los objetivos se definen en función del logro de los estándares.

Aseveraciones relativas a los estándares tenderían a expresarse en frases del tipo siguiente: «dentro del tema de la producción de textos, los estudiantes de cuarto grado deben estar en condiciones de comprender el siguiente conjunto de conceptos claves propios de la lengua». Aseveraciones relativas a objetivos se expresarían en frases del siguiente tipo: «aplicar normas de escritura en la producción de textos». O bien, estándar: «los estudiantes de tercer grado deberían apreciar el lenguaje como un medio de comunicación»; objetivo: «redactar una carta o crear un cuento o escribir un texto coherente».

El segundo tipo de estándar se puede traducir en acciones que demuestren niveles de logro y pueden ser traducidos en la expresión numérica de un nivel alcanzado o deseado dentro de un área del conocimiento o del hacer.

Estándar mínimo o básico: «los estudiantes de cuarto grado deben estar en condiciones de lograr una cierta comprensión de los conceptos expresados en el texto y pueden devolver una información parcial del mismo».

Frases relativas a un estándar de nivel múltiple: nivel 1, «el estudiante está en condiciones de lograr una cierta comprensión de los conceptos expresados en el texto y puede devolver una información parcial del mismo»; nivel 2, «el estudiante comprende los conceptos expresados en el texto y puede hacer un resumen de la información»; nivel 3, «el estudiante comprende los conceptos expresados en el texto, puede formular hipótesis al respecto y sacar conclusiones».

Dentro del marco del Laboratorio se prestará atención a ambos tipos de estándares. De gran importancia será el tratamiento de los estándares de contenido. Sin embargo, estándares de contenido sin referencia a los estándares de competencias no tienen mucho sentido. Los primeros indican lo que se debe aprender y hacer, pero dicho conocimiento es de poca utilidad si no se tiene información acerca de cuánto se aprende o se puede hacer a partir de ellos. Los juicios a partir de la medición servirán para que cada país pueda situarse en relación al estándar.

250

La identificación y formulación de estándares se hará en primer lugar a partir de las formulaciones curriculares de los países. Sin embargo, también se tomarán en cuenta otros criterios tales como el estado de avance de la disciplina; la relevancia social de ciertas competencias en relación a su utilidad para la vida, y estándares internacionalmente reconocidos en estudios nacionales e internacionales.

10. Características de las variables de producto a ser medidas

El tipo de características cognoscitivas susceptibles de desarrollarse por la vía del aprendizaje más o menos formal, aunque constituye un tema eminentemente académico y por ello no directamente ligado al propósito del Laboratorio, el cual está eminentemente relacionado con aspectos de la praxis de la Educación, no es posible, sin embargo, soslayarlo en este planteamiento. Desde el concepto de Inteligencia General, designado más técnicamente como factor «g», hasta nuestros días, la tendencia ha oscilado hacia considerar cada vez mayor el ámbito de los aprendizajes opuesto al de las aptitudes. Sin embargo, el tema está lejos de dilucidarse.

En esta sección se pretende dar una visión del objeto de que se trata, en el Laboratorio, cuando se habla de apreciar, medir o evaluar variables

de producto. El origen de la necesidad de arribar a una especificación nace de una observación de la mayoría de los procesos que ocurren en la casi totalidad de los sistemas escolares. Tal observación arroja como conclusión que, en casi todos los casos, el énfasis de la actividad educacional formal o escolar está puesto más en llevar al alumno hacia el *conocer* que hacia el *hacer*.

Diversas formas de integrar las aptitudes y los productos del aprendizaje se han intentado.

Una de ellas consiste en establecer un esquema que ubica a las distintas variables cognoscitivas propias del ser humano en categorías, en cada una de las cuales se produce una combinación de condiciones, a saber: la generalidad del rasgo y la modificabilidad del mismo. De alguna manera se puede sugerir que ambas podrían visualizarse como sumando algebraicamente 1. Mientras mayor sea la amplitud de la aplicabilidad del rasgo a situaciones variadas, ello es la generalización, menor será la posibilidad de modificación del rasgo y viceversa. Podría, pues, concluirse que mientras mayor es el rasgo de aplicación de una característica cognoscitiva, menor debería ser su posibilidad de alterarse a través del aprendizaje. Si se intenta una clasificación, la más simple, se obtienen dos niveles: Inteligencia General, de amplia generalización y de casi nula modificabilidad y los aprendizajes concretos, propios de asignaturas escolares, de reducida generalización y de altísima modificabilidad. Para acercarse a un cuadro más preciso, es posible agregar un nivel intermedio, ello es de aplicación medianamente general y de modificabilidad posible, aunque condicionada por la duración del proceso de modificación, aquí es posible ubicar a la Comprensión Lectora en el ámbito verbal y a la Habilidad Lógico-Matemática en el campo cuantitativo.

Otro enfoque consiste en construir una jerarquía de todo aquello posible de aprender, pero que de algún modo implica en sus niveles superiores a rasgos que se acercan cada vez más a las aptitudes y ello por dos razones. La primera, por cuanto se alejan cada vez más de lo meramente memorizable. La segunda razón, por cuanto este esquema jerárquico o taxonómico, construye cada categoría sobre la anterior y en ese sentido las superiores alcanzan niveles de generalización amplísimos. La más conocida clasificación involucra seis niveles: conocer (memorístico), comprender (que involucra diferencias en la expresión de lo aprendido con respecto a cómo fue aprendido), aplicar (usar el conocimiento en situaciones nuevas), analizar (descomponer orgánicamente el todo en sus partes constitutivas), sintetizar (producir un todo organizado y único), evaluar (contrastar el valor de algo con respecto de patrones intrínsecos o extrínsecos).

Una tercera posibilidad está basada en una jerarquía diferente y más atingente al tema del conocer y del hacer y que está basada en la medida en que un aprendizaje es transferible. Transferible desde respuestas específicas hasta discriminaciones, luego a clasificaciones, que a su vez se transfieren a reglas y ellas hacia comportamientos regidos por reglas tales como la resolución de problemas.

Una cuarta alternativa lleva a un planteamiento más amplio que integra rasgos psicológicos variados y que se refiere a dominios del aprendizaje. Entre ellos se ubican las habilidades motoras, la información verbal, las habilidades intelectuales, las estrategias cognoscitivas, y, finalmente, las actitudes.

Otra forma de abordar el tema lleva a considerar que el aprendizaje dice relación, en última instancia, con dos actividades fundamentales: resolver situaciones problemáticas y generar tales situaciones. Los problemas, a su vez, se sitúan en siete campos, a saber: lingüístico, musical, lógico-matemático, espacial, corporal-kinestésico, intrapersonal e interpersonal.

Todo lo anterior conduce, al menos, a cuatro conclusiones.

- La primera se refiere a que hay aprendizajes de tipo más generalizable y otros más específicos.

- La segunda se refiere a que los aprendizajes se sitúan en una relación jerárquica.

- La tercera indica que una proporción significativa de lo que el hombre puede conocer o hacer, surge de aprendizajes más o menos formales.

- La cuarta se refiere a que no siempre es válido establecer diferencias dicotómicas entre aprendizajes exclusivamente memorísticos o aprendizajes de mayor aplicabilidad.

Todo lo anterior lleva a sugerir que una medición y evaluación de los aprendizajes deba dirigirse a un espectro o continuo de ellos.

11. Diseño de un sistema de medición internacional y regional

Antes de plantear la idea de un sistema de medición que agrupe a varios países, es conveniente establecer las diferencias que existen entre un sistema de evaluación regional o nacional y los exámenes aplicados en forma regular en cada sistema escolar. La razón para esto

radica en el hecho de que la expresión más frecuente de la evaluación que se conoce a nivel masivo en los países del área está constituida precisamente por tales exámenes.

DIFERENCIAS ENTRE LAS EXAMINACIONES Y LAS EVALUACIONES NACIONALES

	EXAMINACIÓN	EVALUACIONES NACIONALES O REGIONALES
SUJETOS	INDIVIDUOS	MUESTRA DE INDIVIDUOS
UNIDAD DE MEDICIÓN	INDIVIDUOS	ESCUELA, DISTRITO, PROVINCIA, PAÍS
FRECUENCIA	ANUAL	BI O TRIENAL, POR TEMA
CONTENIDO	CURRÍCULUM COMPLETO	MATEMÁTICA, CIENCIA, LENGUAJE ESCRITO Y ORAL
MOMENTO	FIN DE CICLO	UN GRADO POR CICLO
ENFOQUE DE MEDICIÓN	NORMAS O CRITERIOS	CRITERIOS

12. Tipo de análisis: transversal o longitudinal

Una temática como la atingente al Laboratorio tiene dos posibilidades básicas de análisis: transversal o longitudinal. Los análisis evaluativos transversales corresponden a la evaluación de los alumnos que se hace en un momento de su vida escolar. Los análisis evaluativos longitudinales son aquellos en que la evaluación de los mismos alumnos se hace al menos en dos momentos de su trayectoria escolar.

Un estudio transversal se enfoca en el aprendizaje acumulativo hasta un cierto grado de escolaridad o hasta una edad determinada. Un estudio de medición longitudinal se enfoca en las características instruccionales que dan origen al aprendizaje dentro del período comprendido entre los momentos de medición.

Es importante señalar, eso sí, que el hecho de que una medición transversal se realice periódicamente, y con cohortes distintas de alumnos, no la hace equivalente a una medición longitudinal, la que requiere, por definición, trabajar con una misma cohorte, pero en distintos momentos.

La mayoría de los estudios llevados a cabo por entes internacionales son transversales, dado que tienen por objetivo evaluar y comparar el rendimiento acumulativo en un área de materias, al llegar a un grado o en una edad determinados. Son además de este tipo por el costo y la dificultad que implica el seguimiento de los individuos de la cohorte en estudio, propio de los análisis longitudinales. El campo de contenidos se ha definido como el núcleo esencial del área de materias acumuladas que culmina en el grado o año fijado como objetivo. Se trata de saber en detalle cuánto saben los alumnos y qué son capaces de hacer. Las variables explicativas pueden ser individuales (qué tipos de alumnos tienen tal o cual rendimiento) o sistémicas (qué escuelas y cuál contexto pedagógico se relacionan con un mayor rendimiento).

El Laboratorio ha optado por un análisis transversal, al menos en una primera etapa. La idea es evaluar el aprendizaje acumulado hasta el grado de tercer o cuarto año de Primaria, nivel en el que, en los países de la muestra, se completa el primer ciclo. Lo anterior se basa en la concepción de que los efectos de la Educación son eminentemente acumulativos y jerárquicos, por lo cual al evaluar en un cierto nivel, se está también haciéndolo con respecto a los niveles anteriores.

La acción del Laboratorio, como una forma de tomar en cuenta el efecto de la variación que existe entre los países, consideró aplicar cuestionarios sobre la estructura de los programas de estudios anteriores al grado objetivo. Lo anterior, aunque se basó en la condición acumulativa y jerárquica del aprendizaje escolar ya comentada.

13. Enfoque de medición: referido a normas o a criterios

Las mediciones pueden estar referidas a normas o a criterios, no existiendo connotación valórica o cualitativa alguna asociada a cada enfoque. Lo anterior implica que el empleo de un enfoque obedece al propósito que se busca con la medición. La que sí es insoslayable es la necesaria congruencia entre enfoque y propósito.

La diferenciación más simple entre ambos enfoques implica que en los instrumentos referidos a normas, el intento es comparar al estudiante, en una variable determinada, con la norma de su grupo. Las mediciones normativas no entregan información respecto de cuánto saben y cuánto no saben los sujetos. Ellas reportan cuál es la posición relativa de un sujeto con respecto de la norma de su grupo. Este tipo de instrumento entrega al maestro y a quienes diseñan las políticas, información muy limitada que pueda ser útil para mejorar las escuelas.

Los instrumentos referidos a criterios están enfocados a determinar si un individuo ha logrado un nivel predeterminado o criterio prefijado, en una variable dada. Tales instrumentos, que requieren descripciones detalladas de qué miden, suponen entregar información precisa sobre logros respecto de un estándar absoluto. Los niveles corresponden a estándares, metas o porcentajes de logros. En algunos casos, para determinar dichos niveles es preciso considerar variables afectivas, en cuanto a prevenir consecuencias negativas en ese dominio, a causa de un exageradamente alto nivel de logro en lo cognoscitivo.

En la escuela y más allá de ella, la información provista por tales instrumentos apoya, de preferencia, a maestros que desean mejorar su instrucción, a diseñadores de políticas que quieren conocer cómo mejorar las escuelas, a directores e inspectores que necesiten apoyar a los profesores, a los diseñadores de currículum y de textos que requieran verificar la medida en que el currículum y los textos son efectivos.

Existe una variedad de formas de especificar la temática incluida en cada ítem de un instrumento referido a criterios. Tales formas se sitúan en un continuo, desde un mínimo hasta un máximo de especificidad.

En el extremo más carente se sitúan los ampliamente conocidos **objetivos en términos de conducta**. A continuación se ubican los llamados **formatos de ítems**, que consisten en un conjunto de reglas altamente detalladas que debieran derivar en ítems de naturaleza muy homogénea. Siguen los conocidos como **objetivos ampliados**, que no son más que versiones más elaboradas y detalladas de los objetivos en términos de conducta. Seguidamente se encuentra una **estrategia** conocida como **de enfoque limitado**, que es básicamente una focalización en pocos objetivos de naturaleza terminal, que incluyen a otros objetivos en tránsito. Finalmente, se ubica el **método** denominado **de especificación de instrumentos** que se deriva de los formatos de ítems antes mencionados.

Las especificaciones en un formato ortodoxo deben incluir cinco componentes. El primero consiste en una descripción general del instrumento, limitado a un par de líneas.

El segundo consiste en uno o varios ítems de ejemplo, que incluyan las instrucciones de respuesta para el estudiante.

El tercero corresponde a una descripción de los atributos del estímulo, que se refieren fundamentalmente a las limitaciones a que deben ceñirse las preguntas o ítems del instrumento.

El cuarto componente corresponde a los atributos de respuesta. Estos varían de acuerdo al tipo de ítems de que se trate: de selección o

de respuesta abierta. Lo que se pretende, en este caso, es no sólo indicar la naturaleza de las respuestas correctas, sino también la de todas las respuestas erróneas posibles de anticipar. En el caso de ítems abiertos, se trata de establecer criterios que permitan determinar la bondad de una respuesta.

Finalmente, el quinto componente se denomina suplemento de especificaciones y se relaciona con toda aquella información adicional que se requiera para la construcción del instrumento.

Las especificaciones permiten aumentar la validez de contenido de un instrumento. Constituyen una guía para los redactores de ítems, en cuanto a que todos escriban ítems de similar validez y nivel de dificultad. Versiones simplificadas de las especificaciones son usadas para informar a profesores y estudiantes acerca del instrumento que se aplicará. Permiten la creación de bancos de ítems para proveer ítems en diferentes situaciones y tiempos, mientras que aseguran razonablemente la consistencia interna del instrumento como también su validez y aproximadamente un mismo nivel de dificultad.

Existe un aspecto adicional importante de señalar y se relaciona con que en los instrumentos referidos a criterios debe tenerse especial conciencia que las medidas de calidad (validez y confiabilidad) que se basan en el uso de coeficientes de correlación, se ven afectadas por la posible mayor homogeneidad de los resultados. Lo anterior, dado que esa homogeneidad dará origen a una menor varianza y, por ende, a subestimaciones de los índices.

256

A P É N D I C E 1

Elaboración de los instrumentos de medición

Los cuestionarios y pruebas que se utilicen en el Laboratorio serán contruidos de acuerdo a los procedimientos internacionalmente aceptados. Tales procedimientos incluyen, al menos, las siguientes etapas:

- a. Definición de los objetivos del instrumento
- b. Selección del tipo de preguntas por objetivo
- c. Construcción de los ítems o reactivos
- d. Edición de los ítems o reactivos
- e. Ensamblaje de los instrumentos
- f. Evaluación formativa del instrumento completo por medio de juicios de expertos

- g. Validación de los instrumentos en cada país
- h. Adaptación de los instrumentos al léxico de cada país
- i. Determinar la reacción de los estudiantes a los instrumentos
 - al tipo de preguntas
 - al tiempo total para responder las pruebas
 - conocer tasa de omitidas
 - conocer la tasa de adivinación
 - conocer el nivel general de dificultad
- j. Análisis psicométrico de instrumentos
 - funcionamiento de las alternativas
 - elaboración de la lista de posibles respuestas a los ítems abiertos
 - formación de un banco de productos largos
 - determinación de la interacción ítem-escuela e ítem-país
 - conocimiento de la reacción de los profesores a los ítems y pruebas
 - conocimiento de la reacción de estudiantes, profesores, directivos y padres a los cuestionarios
 - determinación final de ajustes a los instrumentos
- k. Forma definitiva de los instrumentos.
- l. Verificación de procedimientos de campo.
 - entrenamiento de aplicadores
 - procedimiento de aplicación de pruebas y cuestionarios

Metodología general

Para la mayor parte de los instrumentos se optó por desarrollar primariamente ítems de respuesta cerrada. La decisión se basó en el gran tamaño de la muestra, el tiempo disponible y la necesidad de aplicar las pruebas en forma masiva.

Construcción de pruebas de lenguaje y matemática

1. Selección de los objetivos

La medición se basó en las preferencias curriculares seleccionadas por cada coordinador de país, al cual se le solicitó que informara en relación a un listado de contenidos a cuál curso o grado correspondía el

tratamiento inicial del contenido y en cuáles cursos se enfatizaba. Con esa información se construyó una tabla de especificaciones, que fue entregada a los constructores de los instrumentos. Complementariamente, se revisaron los objetivos de lenguaje y matemática para tercer y cuarto año básico de cada país.

Criterios usados para la selección de los objetivos

- a. Grado de importancia asignado en cada país,
- b. Grado en que la conducta es observable en forma masiva.

2. *Construcción de ítems y reactivos*

a. Participación de algunos países, aportando ítems. Se solicitó a todos los países participantes en el Laboratorio que enviaran al menos 10 ítems de lenguaje y de matemática, para ser incorporados a las pruebas y comenzar así a construir el banco regional de ítems.

b. Elaboración de ítems por cada objetivo seleccionado. Se redactó una cantidad mayor de ítems que los requeridos, con la finalidad de seleccionar los que mostraran una mayor validez de contenido. Los ítems, con excepción de la prueba de redacción, son de selección múltiple con cuatro alternativas.

c. Análisis cualitativo de cada ítem, por juicio de expertos. El análisis se basó en:

- la correspondencia entre ítem y objetivo
- la coincidencia en cuanto a cuál es la clave de respuesta
- lo adecuado de los distractores presentados
- la correcta redacción de enunciado y las alternativas
- la inteligibilidad del lenguaje utilizado para el nivel de desarrollo cognoscitivo de los examinados.

d. Aprobación de los ítems seleccionados por parte del grupo de profesores y especialistas.

e. Clasificación de los ítems según grado de dificultad estimado por los profesores de acuerdo con su experiencia, en tres categorías: fáciles, de mediana dificultad o difíciles.

3. *Ensamblaje de los instrumentos*

Se construyeron formas equivalentes de cada instrumento por cada nivel. En el ensamblaje se procuró mantener la equivalencia entre las

diferentes formas en cuanto a objetivos cubiertos, así como en el grado de dificultad estimado.

APÉNDICE 2

Población y muestra para la aplicación experimental de las pruebas de lenguaje y matemática

Primeramente, es imprescindible tener presente cuál es el propósito de la aplicación experimental de las pruebas de Lenguaje y Matemática. Se trata de obtener información respecto de, al menos, seis aspectos, a saber: grado de adecuación de las pruebas a la realidad educacional de cada país, incluyendo las variaciones en el uso del lenguaje; reacción de los estudiantes con respecto de las pruebas; antecedentes para un primer análisis de ítems; reacción de los profesores respecto de las pruebas; retroinformación que permita los ajustes necesarios a la pruebas; antecedentes de la aplicación de los procedimientos de campo.

Es igualmente importante mantener en mente que esta aplicación no tiene por objetivo establecer aún conclusiones respecto de los aprendizajes en Lenguaje y Matemática, y que incluso es altamente conveniente resistir la inclinación de usar los datos para propósitos que no sean los de validación de instrumentos y análisis de ítems.

La unidad de análisis es el alumno del último grado del primer ciclo de Educación Básica. Tal grado es, en algunos países, el Tercero y en otros el Cuarto. Por otra parte, por razones de calendarios escolares y estructuras curriculares, parece conveniente incluir ambos en grados en la muestra para la aplicación experimental.

El universo está constituido por todos los alumnos que cumplen con las características propias de la unidad de análisis. Por razones prácticas, tal universo ha sido acotado a lo que se denominará población muestral. La ventaja de mantener presente cuál es el universo reside en que las exclusiones deben ser explícitamente justificadas, permitiendo así apreciar la magnitud y consecuencias de tales exclusiones.

El diseño muestral debe especificar los criterios de muestreo, los cuales en el presente caso se definen como aquellos propios de una muestra estratificada. Lo anterior implica que los procedimientos de muestreo no son exclusivamente aleatorios, con las consecuencias de aumento del sesgo y subjetividad consiguientes.

La intención es estratificar por tipos de escuela, considerando cuatro tipos, a saber: escuelas privadas de elite; escuelas públicas de

nivel socioeconómico medio de ciudades grandes, escuelas públicas de nivel socioeconómico bajo de ciudades grandes y escuelas rurales.

El procedimiento contempla la elaboración de una lista de escuelas por país para cada uno de los estratos especificados anteriormente. Las listas constituirán el espacio muestral del cual seleccionar aquellas que serán incluidas en la aplicación experimental.

Se seleccionarán las escuelas «por juicio» (deliberadamente), por estrato, de acuerdo al siguiente esquema:

ESCUELAS DE ELITE:	2 por país
ESCUELAS PÚBLICAS DE CIUDADES GRANDES, DE NIVEL SOCIOECONÓMICO MEDIO:	3 por país
ESCUELAS PÚBLICAS DE CIUDADES GRANDES, DE NIVEL SOCIOECONÓMICO BAJO:	3 por país
ESCUELAS RURALES:	3 por país
TOTAL:	11 por país

260

El diseño considera seleccionar un curso de Tercer Grado y uno de Cuarto Grado por escuela. Ambos deberán ser seleccionados al azar, salvo que exista uno solo de cada grado en la escuela respectiva.

Se considera en cada curso tomar la totalidad de los alumnos. Se ha estimado un tamaño promedio de los cursos de 35 alumnos.

Lo señalado arroja una muestra por país de 770 niños y un total para la aplicación de 1078 estudiantes.

Es importante señalar que los criterios de estratificación propuestos deben ser necesariamente acordados con los países, dadas la amplia variabilidad de situaciones existentes y las diferentes conformaciones de los sistemas de información estadística educacional en cada caso.

Las precisiones anteriores dan origen a tres interrogantes, las cuales se exponen a continuación:

1. ¿Es conveniente incluir o excluir las escuelas unidocentes?
2. ¿Es preciso incorporar el tamaño de la escuela como criterio de estratificación?
3. ¿Es necesario considerar la condición de coeducacional de la escuela como criterio de estratificación?

APÉNDICE 3

Consideraciones respecto del control de calidad de las acciones a emprender

Una de las características de las pruebas que es fundamental para el éxito de las acciones del Laboratorio es la relevancia o pertinencia de los instrumentos. Por lo anterior, una de las etapas que se ha considerado es la revisión por parte de los países de las matrices curriculares que dieron origen a las pruebas y su posterior contrastación con cada una de ellas.

Un segundo aspecto se relaciona con la necesidad de que los instrumentos respondan a las diferencias de lenguaje propias de los diferentes países. Esto varía desde la necesaria traducción para el caso del Brasil hasta la adaptación a las diferencias lingüísticas entre los países de una misma lengua. Para lo anterior se desarrollará una verificación con cada país al respecto. Del mismo modo se procurará detectar y evitar los sesgos culturales que pudiesen manifestarse.

El hecho de que las mediciones se lleven a cabo en entornos tan diferentes, refuerza la necesidad de que se efectúe un proceso de validación de los ítems y de las pruebas, el cual ya fue descrito previamente.

La aplicación de los instrumentos requiere, para producir resultados comparables, el que sea abordado en forma similar en todos los casos. Por lo anterior, se construirán los manuales de procedimientos respectivos y se pondrán a disposición de los países las acciones de entrenamiento que se puedan requerir.

El manejo de la variedad y cantidad de información que se generará implica la construcción de un sistema de almacenaje y acceso a los datos. En efecto, se ha estructurado tal sistema, el cual será ampliamente informado a los países y se ofrecerán asimismo las acciones necesarias de entrenamiento al respecto.

Se pondrá en marcha, en todos los casos, una supervisión independiente y externa de las aplicaciones de los instrumentos que asegure ortodoxia y comparabilidad de resultados.

Existirá en todos los casos un control de calidad externo al sistema, el que permita otorgar a las acciones un nivel de calidad internacional.