



2009 - Volumen 2, Número 1



RINACE

<http://www.rinace.net/riee/numeros/vol2-num1.html>

Revista Iberoamericana de Evaluación Educativa

La Revista Iberoamericana de Evaluación Educativa es una publicación oficial de la Red Iberoamericana de Investigación sobre Cambio y Eficacia Escolar.

Dirección:

F. Javier Murillo
Marcela Román

Coordinación Editorial:

Verónica González

Comité Directivo:

Marcela Gajardo (PREAL - Chile)
Sergio Martinic V. (PUC - Chile)
Carlos Pardo (ICFES - Colombia)
Margarita Poggi (IIPE/UNESCO - Argentina)
Francisco Soares (UMG - Brasil)
Héctor Valdés (ICCP - Cuba)

Comité Científico:

Juan Manuel Álvarez Méndez (UCM - España)
Patricia Arregui (GRADE - Perú)
Daniel Bogoya (UPN - Colombia)
Nigel Brooke (UMG - Brasil)
Leonor Cariola (MINEDUC - Chile)
María do Carmo Clímaco (Portugal)
Cristian Cox (PUC - Chile)
Santiago Cueto (Grade - Perú)
Tabaré Fernández (ANEP - Uruguay)
Juan Enrique Froemel (Min. Ed. – Qatar)
Reyes Hernández Castilla (UAM – Castilla)
Rubén Klein (Fund. Cesgranrio – Brasil)
Luis Lizasoain (UPV - España)
Jorge Manzi (PUC - Chile)
Felipe Martínez Rizo (INEE - México)
Joan Mateo (UB - España)
Liliana Miranda (MINEDU - Perú)
Carlos Muñoz-Izquierdo (UIA – México)
Margarita Peña (ICFES - Colombia)
Dagmar Raczynski (Asesorías para el desarrollo - Chile)
Héctor Rizo (UAO - Colombia)
Guadalupe Ruiz (INEE - México)
Mario Rueda (UNAM - México)
Ernesto Schiefelbein (CIDE – Chile)
Alejandra Schulmeyer (INEP - Brasil)
Javier Tejedor (USAL – España)
Flavia Terigi (UBA/UNGS - Argentina)
Alexandre Ventura (IGE - Portugal)

© Derechos Reservados RINACE 2009



Editorial

- La Evaluación de los Aprendizajes Escolares: un Recurso Estratégico para Mejorar la Calidad Educativa** 4
Marcela Román y F. Javier Murillo

Sección Temática: Evaluación de Resultados Escolares

- La Efectividad y la Eficacia de las Mediciones Estandarizadas y de las Evaluaciones en Educación** 10
Juan Enrique Froemel
- La Experiencia Chilena de Instalación de Estándares de Aprendizaje y la Nueva Institucionalidad para la Calidad de la Educación** 29
Pedro Montt
- La Experiencia del Sistema Nacional de Evaluación del Proceso Educativo, SNEPE, 48 en Paraguay. Aprendizajes y desafíos**
Marta Lafuente
- La Calidad del Sistema Educativo Mexicano desde los Resultados de Evaluaciones Nacionales. El aprendizaje en Matemáticas** 74
Guadalupe Ruíz
- Honduras 2008: La Evaluación de los Aprendizajes Vislumbra una Luz de Esperanza en Educación** 90
Russbel Hernández
- La Evaluación de la Educación Cívica y Ética en México. Un Recorrido por 3° de Primaria y 3° de Secundaria** 114
Oscar Barrera
- La Evaluación de la Escritura para la Evaluación de la Lectura** 130
Fabio Jurado
- La Evaluación de Aprendizajes de Personas Jóvenes y Adultas** 147
María Eugenia Letelier
- El Análisis Cognitivo de Tareas, Base para el Diseño de Instrumentos de Evaluación en el Aprendizaje en Línea** 162
Eduardo Peñalosa y Sandra Castañeda
- La Evaluación del Logro Escolar en la Educación Básica en México. Los Exceles de Matemáticas** 186
Andrés Sánchez
- Las Prácticas de los Comités de Evaluación del Aprendizaje Escolar. Algunas Paradojas a Partir del Caso Mexicano** 205
Adriana Andrade

Sección Temática Libre

- Análisis Multinivel de los Factores Asociados a la Educación Cívica en Estudiantes de Secundaria** 226
Joaquín Caso, Eduardo Hernández y Juan Carlos Rodríguez
- Da Avaliação de Professores: Reflexões Sobre o Caso Português** 239
Pedro Montt
- ¿Cómo Evaluar la Convivencia Escolar? La Mirada Desde los Estudiantes** 257
Álvaro Carrasco



LA EVALUACIÓN DE LOS APRENDIZAJES ESCOLARES: UN RECURSO ESTRATÉGICO PARA MEJORAR LA CALIDAD EDUCATIVA

Marcela Román y F. Javier Murillo

Revista Iberoamericana de Evaluación Educativa 2009 - Volumen 2, Número 1

<http://www.rinace.net/riee/numeros/vol2-num1/editorial.pdf>



En el principio y origen de la evaluación de carácter más claramente educativo, se encuentran los esfuerzos de maestros y maestras por determinar los aprendizajes adquiridos por sus estudiantes de manera de poder orientar la enseñanza y el desarrollo de ulteriores actividades educativas. Así, la evaluación de los aprendizajes y resultados que alcanzan los niños, niñas y los jóvenes en su proceso educativo es sin duda el área que aporta el mayor acumulado de experiencia y conocimiento a la evaluación en el campo educativo. En la actualidad, dicha práctica ha sobrepasado los límites del aula y la escuela para constituirse en la base de las evaluaciones nacionales e internacionales del rendimiento; en el primer caso prácticamente como indicador exclusivo de la calidad de la educación ofrecida en cada país, y en el segundo como forma de comparar cuánto saben y rinden los estudiantes en los distintos países del mundo y, con ello la calidad de los propios sistemas.

Sin embargo, a pesar del importante conocimiento acumulado, los aportes de la evaluación al fortalecimiento de los aprendizajes y a la calidad educativa de los sistemas están aún muy lejos de lo necesario, y muy lejos también de las expectativas cifradas en esta práctica, cuya potencialidad va más allá del simple hecho de medir el desempeño escolar o dar cuenta del avance, estancamiento o retroceso en los aprendizajes que adquieren los estudiantes en los distintos los países y en la región. Más que evaluar para aprender y comprender cómo y porqué aprenden lo que aprenden niños, jóvenes y adultos en su proceso educativo, la evaluación de los aprendizajes parece haberse conformado con medir cuanto rinden o logran los estudiantes de acá o allá, en Lengua, Matemática y algunas otras disciplinas que hacen parte del currículo escolar.

Y esta mirada del todo fragmentada y limitante de las reales posibilidades de la evaluación nos acompaña no sólo desde las evaluaciones estandarizadas, sino que también se observa y ratifica como práctica mayoritaria en los profesores y profesoras a la hora de evaluar los aprendizajes de los estudiantes. La evidencia mostrada por estudios y experiencias de apoyo al mejoramiento de procesos de enseñanza y aprendizaje en escuelas latinoamericanas, sobre todo en contextos de exclusión y vulnerabilidad ratifica la idea de que con la evaluación sólo se busca conocer cuánto saben los estudiantes para así calificarlos y ubicarlos entre los que eventualmente aprobarán o reprobarán, o quienes están por debajo o sobre el promedio.

Débil también ha sido el aporte de la evaluación de los aprendizajes a la disminución de la inequidad que se reproduce en los sistemas educativos, y que se replica y ratifica en cada escuela y al interior del aula. En efecto, a estas alturas y aún a riesgo de parecer pesados, es necesario volver a insistir en que no es posible mejorar la calidad educativa sin reducir las importantes brechas entre los estudiantes al interior de las escuelas y los sistemas. Dicho en otras palabras, la inequidad educativa pone techo y límite a la posibilidad de incrementar la calidad y sin duda la evaluación tiene mucho que aportar y decir respecto de cómo lograr aprendizajes significativos y estables para todos, independientemente del género, pertenencia o nivel socioeconómico del estudiante, del contexto o ubicación de la escuela. Sirvan dos afirmaciones más para reforzar lo anteriormente manifestado: las escuelas y los sistemas requieren evaluar los aprendizajes que adquieren los estudiantes para conocer y dar cuenta de su calidad y niveles de avance. Sin embargo, la evaluación de los aprendizajes por sí sola no puede mejorar la calidad de lo aprendido, por muy sistemática y rigurosa que ella sea.

Pero no caigamos en el pesimismo: desde la evaluación es posible –y necesario– generar información relevante sobre factores y variables que limitan o posibilitan los aprendizajes de los niños y los jóvenes de manera de poder actuar oportunamente y en consecuencia. La evaluación puede incidir también en procesos educativos más integrales y acordes a las necesidades actuales. En tal sentido, hay al menos tres exigencias que hacerle a la evaluación de los aprendizajes y resultados escolares:

- i. La evaluación debe validar y reforzar la idea de que la educación busca el desarrollo integral de la persona, dado que ella define lo que se hace y se prioriza en educación. Sabiendo entonces que lo que es evaluado pasa a ser la prioridad en las aulas, las escuelas y los sistemas, convirtamos a la evaluación en un ejemplo. De esta forma las evaluaciones nacionales e internacionales han de avanzar en la incorporación de otras áreas disciplinarias y grados que favorezcan una intencionalidad pedagógica y educativa que se haga cargo y responda por el desarrollo integral y armónico de todos los ciudadanos.
- ii. La evaluación necesita contextualizar los aprendizajes medidos. Es tiempo que el alto nivel de desarrollo tecnológico y metodológico que ha alcanzado la medición de los aprendizajes escolares, vaya de la mano con la identificación y comprensión de las condiciones y factores intra y extra escuela que favorecen y hacen posible que los estudiantes alcancen aprendizajes significativos, relevantes y estables.
- iii. La evaluación ha de ofrecer criterios y estrategias para mejorar su uso a nivel de las escuelas y las comunidades educativas. En efecto, todo lo anterior pierde su norte y sentido si es comprendido y utilizado sólo por los técnicos, autoridades educativas o los investigadores. Al respecto vale la pena destacar el rol privilegiado y la responsabilidad que le cabe a la evaluación en devolver a las escuelas y sus actores (docentes, directivos, padres y estudiantes), los resultados de tales procesos de una manera cercana y útil que les permita comprender los logros como consecuencia de condiciones, factores, decisiones y acciones que ocurren cotidianamente y convergen en la escuela. La evaluación bien comprendida y utilizada es un aliado estratégico e insustituible para mejorar los procesos de enseñanza y aprendizaje y con ello, el desempeño o rendimiento escolar.

Este número de la revista, que coincide con el primer aniversario de su aparición, recorre la práctica de evaluación de los aprendizajes desde la sala de clases hasta los sistemas educativos. A través de dicho recorrido y de la mano de varios ex Vice-Ministros de Educación, coordinadores de unidades, centros y sistemas nacionales de evaluación educativa, académicos e investigadores, es posible conocer los esfuerzos, tensiones y desafíos detrás de la implementación de evaluaciones externas y estandarizadas, sus aportes o pendientes a la calidad y equidad educativa. Al mismo tiempo ofrece la ocasión de profundizar no sólo en los resultados de evaluaciones realizadas, sino que participar de sus estrategias, procesos e instrumentos en niveles y áreas tan diversas como la escritura, la ciudadanía, la convivencia y por cierto en aquellas históricas como Lengua y Matemática tanto en la educación primaria y secundaria como en jóvenes y adultos fuera de los sistemas educativos formales.

Abre la revista el texto de Enrique Froemel, ex coordinador del *Laboratorio Latinoamericano de Evaluación de Calidad de la Educación* de la UNESCO y responsable de la puesta en marcha del sistema nacional de evaluación del rendimiento educacional de Qatar. En dicho texto, se discute la efectividad y eficacia de la medición y evaluación estandarizada, teniendo como telón de fondo los estudios

internacionales de evaluación del rendimiento escolar. Junto con la entrega de elementos políticos, teóricos y metodológicos, analiza empíricamente las limitaciones propias de las mediciones estandarizadas y propone cuales han de ser las condiciones para que ellas se constituyan en reales herramientas para una política educativa que busca mejorar la calidad de los sistemas.

Pedro Montt, desde su privilegiada posición como coordinador general de la *Unidad de Curriculum y Evaluación* del Ministerio de Educación de Chile y ex Subsecretario de Educación, comparte las perspectivas, implicancias y consecuencias del proceso de elaboración de estándares de aprendizaje que ha buscado impactar en la calidad y equidad del sistema educativo chileno. Así y en el marco de la nueva institucionalidad que se desarrolla en el país, presenta y discute ventajas, tensiones y desafíos de la instalación de *Mapas de Progreso* y *Niveles de logro*, en tanto estándares de contenido y desempeño que asumen el aprendizaje como un continuo que se enriquece a lo largo de la trayectoria escolar.

El artículo de Marta Lafuente, Vice-Ministra de Educación del Paraguay entre los años 2003 y 2008, recorre la experiencia de la instalación y desarrollo del *Sistema Nacional de Evaluación del Proceso Educativo, SNEPE*, en el marco del proceso político y de Reforma Educativa de los años noventa. Desde allí revisa la situación educativa del país, los aportes, desafíos y temas pendientes del Sistema de Evaluación en tanto una de las políticas estratégicas de la Reforma Educativa. Levanta hipótesis para su desarrollo futuro en el actual momento político del país y desde la nueva gestión del Ministerio de Educación y Cultura.

Guadalupe Ruiz, ex-Directora de Evaluación de Escuelas del *Instituto Nacional para la Evaluación de la Educación de México* (INEE), analiza los resultados de las evaluaciones de la calidad del sistema educativo mexicano realizadas entre 2005 y 2007, deteniéndose especialmente en Matemática en algunos grados de primaria y secundaria. El texto deja en evidencia la enorme desigualdad del desempeño escolar entre las principales modalidades educativas, así como la gran distancia que aún le resta por revertir a la educación mexicana a fin de alcanzar niveles de logro satisfactorios en la mayoría de sus estudiantes. En la misma línea se inscribe el texto de Andrés Sánchez, investigador del INEE de México, que describe las pruebas Excale de Matemáticas y presenta una panorámica del rendimiento académico de los estudiantes en preescolar, algunos grados de primaria y secundaria en dicha área, a partir de las evaluaciones realizadas desde el 2005.

El artículo de Russbel Hernández, Director General de *Evaluación de la Calidad de la Educación de la Secretaría de Educación* de Honduras, pone la nota de esperanza al compartir los avances en los aprendizajes escolares que muestra la reciente evaluación nacional del 2008, luego de más de una década de estancamiento en niveles muy bajos. El autor sostiene que los avances logrados por los estudiantes de primaria en español y matemática, son consecuencia de la introducción de un nuevo currículo para la educación básica, así como de la dotación y distribución de textos escolares y diversos materiales educativos para el trabajo de aula.

Fabio Jurado, Coordinador del *Grupo de Investigación en Evaluación de la Universidad Nacional de Colombia*, en un interesante y provocador artículo, interpela a la escuela en su capacidad para formar lectores y escritores. Desde allí, argumenta y ejemplifica a favor de la evaluación de la escritura para dar cuenta de los niveles de comprensión de lectura, por cuanto el aprendizaje de la lectura presupone el aprendizaje de la escritura. En sus propias palabras: *"La lectura es posible por un diálogo con el texto, lo que supone reconocer que los textos piensan, provocan y seducen hacia la acción"*. De acuerdo con su mirada, la adecuada evaluación de la escritura permite identificar la apropiación semántica de los textos

con los cuales se ha formado quien escribe, pudiendo caracterizar los niveles de dominio en el modo de transmitir un mensaje a un potencial lector.

Los aportes de la escuela y los sistemas escolares a la formación y desarrollo ciudadano de los niños, las niñas y los jóvenes, llegan de la mano de un interesante artículo de Oscar Barrera, coordinador de los *Exámenes de la Calidad y el Logro Educativos (Excale)*, de *Ciencias Sociales* del INEE de México. En dicho texto el autor comparte el proceso y resultados de la evaluación de la educación cívica y ética de los estudiantes mexicanos de primaria y secundaria, profundizando en los desafíos y temas pendientes del Estado en materia de formación de ciudadanos, sobre todo en la educación indígena y rural. Complementa el texto anterior, un sugerente artículo de los docentes e investigadores Joaquín Caso, Eduardo Hernández y Juan Carlos Rodríguez, que analiza y profundiza en los factores asociados a las actitudes cívicas de los estudiantes secundarios en México. A través del modelamiento multinivel, el artículo muestra un efecto reducido de la escuela en la explicación de la conformación de las opiniones y actitudes cívicas en los jóvenes, encontrándose éstas asociadas en mayor medida con las características individuales de los estudiantes.

María Eugenia Letelier, coordinadora del *Sistema de Evaluación y Certificación de Estudios para Personas Jóvenes y Adultas* del Ministerio de Educación de Chile, cambia el foco desde la evaluación de los aprendizajes en el sistema escolar formal, para invitarnos a mirar y reflexionar en torno a la evaluación de aprendizajes en personas jóvenes y adultas. Así y enmarcado en un completo panorama sobre los recientes avances en dicha materia, entrega elementos y argumentos que dejan en evidencia la necesidad de contar con perspectivas y metodologías propias para medir y evaluar los aprendizajes en y para la educación de adultos.

Eduardo Peñalosa y Sandra Castañeda, docentes universitarios e investigadores en Psicología Educativa, entregan su mirada y experiencia en el desarrollo de instrumentos y estrategias para evaluar el aprendizaje de estudiantes de nivel superior en contextos de educación en línea: *el análisis cognitivo de tareas*. Junto con describir y analizar la complejidad de los procesos cognitivos e interactivos implicados en el aprendizaje en espacios y entornos virtuales, se reflexiona y discute en torno a la necesidad de contar con recursos efectivos para evaluar los aprendizajes logrados en las diferentes etapas del proceso de aprendizaje en línea.

El texto de Adriana Andrade, investigadora y docente universitaria, se introduce en la subjetividad e interacción de quienes participan directamente en la construcción de los instrumentos a través de los cuales se evalúa el aprendizaje escolar en México. Así, aborda los *Comités de Evaluación del Aprendizaje Escolar* como espacios sociales que posibilitan el desarrollo de competencias de información, de investigación y de diálogo, contribuyendo a la reconfiguración del espacio educativo y evaluativo como una cuestión de interés público.

El desempeño docente se constituye en el objeto de análisis del artículo de Maria Assunção Flores, académica del *Instituto de Educación de la Universidad de Minho, Portugal*. Durante el recorrido la autora discute perspectivas, supuestos y componentes del modelo de evaluación del desempeño docente que se encuentra en su fase de implementación, en función de sus implicancias y riesgos para la calidad de la enseñanza.

Cierra este número, la invitación de Álvaro Carrasco, investigador chileno asociado al *Centro de Investigación y Desarrollo de la Educación (CIDE)*, para revisar la convivencia escolar. El autor comparte

perspectivas, estrategias y un instrumento que permite evaluarla, surgidos de la experiencia y desde los propios estudiantes.

En definitiva, un seductor número de la Revista Iberoamericana de Evaluación Educativa que esperamos contribuya a una discusión informada que desemboque en la mejora de los sistemas de evaluación de los aprendizajes desarrollados en los diferentes países de la Región; y, con ello, al incremento de la calidad y la equidad de la educación.



LA EFECTIVIDAD Y LA EFICACIA DE LAS MEDICIONES ESTANDARIZADAS Y DE LAS EVALUACIONES EN EDUCACIÓN

Juan Enrique Froemel

Revista Iberoamericana de Evaluación Educativa 2009 - Volumen 2, Número 1

<http://www.rinace.net/riee/numeros/vol2-num1/art1.pdf>

Fecha de recepción: 15 de marzo de 2009

Fecha de comunicación de dictamen: 20 de abril de 2009

Fecha de aceptación: 20 de abril de 2009

1. INTRODUCCIÓN: EL AUGE ACTUAL DE LA MEDICIÓN Y DE LA EVALUACIÓN DEL RENDIMIENTO ESCOLAR

Entre los elementos más difundidos a la vez que debatidos, en la educación formal de niños y jóvenes, en los últimos veinte años, en el mundo, se cuentan la evaluación del aprendizaje escolar y asimismo uno de sus instrumentos fundamentales, la medición estandarizada. Si bien la mayoría de los investigadores y actores en el tema les reconocen su gravitación, no todos están de acuerdo en la real dimensión de su aporte al proceso educativo y lo que es más, variadas percepciones se han construido respecto de ambos, no siempre basadas en bases reales ni menos científicas, generando así mitos y prejuicios.

En este sentido lo que señalara Bloom (1972) acerca de la educación en general, en uno de sus emblemáticos artículos, es aplicable hoy en día a la evaluación y a la medición estandarizada, en cuanto a que *"...seguimos siendo seducidos por el equivalente de los remedios de curanderos, las falsas curas del cáncer, la invención del movimiento perpetuo y las supersticiones..."*. Es así como junto al creciente uso de la medición estandarizada, han surgido legiones de "expertos" en el tema que prometen, por medio de enfoques de medición y evaluación más "humanistas" y "educativos", ir mucho más allá de lo que los métodos basados en la teoría psicométrica son capaces de lograr, desacreditando de paso a estos últimos por ser supuestamente rígidos y estrechos. Lo anterior pese a que existe una masa importante de investigación y de desarrollo teórico, ambos rigurosos, sobre el tema.

El daño real y potencial a la educación y a sus sujetos, los alumnos, es por ello, de tal magnitud, que para colaborar en su prevención, uno de los propósitos centrales de este artículo es intentar precisar y explorar el valor real que, para el proceso educativo formal, tienen la medición y la evaluación de raíz científica. Esto ha implicado centrar la mirada en unos aspectos que no constituyendo una visión exhaustiva, sí tocan aspectos críticos del tema en cuestión. Es así como este trabajo no pretende convertirse en una visión integral acerca de la situación actual de la medición estandarizada en educación, sino que más bien aportar algunas pistas para señalar dónde fijar la mirada.

Aunque mucho se habla hoy en día acerca de medición estandarizada en educación, se ha estimado conveniente precisar aquí, primeramente, qué es lo que hace a un instrumento que sea considerado o no estandarizado.

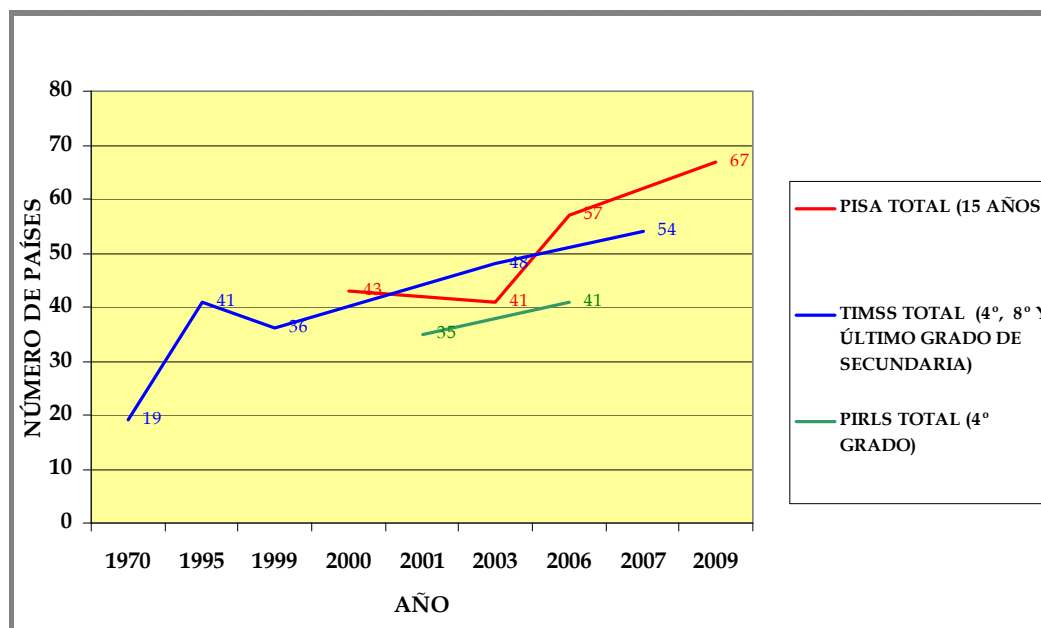
Esta condición implica el cumplimiento de varias características, entre las cuales se cuentan, en primer lugar, el hecho de que los ítems deban ser los mismos y aplicados en un formato equivalente a todos los examinados; en segundo lugar, que las condiciones de administración deban ser también comunes y que, por ende, las instrucciones de aplicación deban ser exactamente a la letra, las mismas, excepto en el caso de las llamadas "acomodaciones" usadas para estudiantes con limitaciones visuales, auditivas o motoras; en tercer lugar, que el procedimiento y proceso de revisión y asignación de puntajes sea también uniforme, no existiendo posibilidad de discrecionalidad en los criterios a aplicar; y finalmente, que la forma en que los resultados son informados también haya de ser común y uniforme. Como es posible apreciar, las condiciones que hacen a un instrumento estandarizado, por una parte, proveen una condición necesaria, aunque no suficiente, para que los resultados sean comparables y por la otra, excluyen de la pertenencia a esta categoría a muchos instrumentos, como es el caso de las pruebas construidas por el profesor o hacen bastante más difícil la estructuración de otros, como ser los *portafolios* si se espera que sus resultados sean comparables.

Una buena forma de visualizar la gravitación de la evaluación y la medición estandarizada es refiriéndose a los estudios internacionales del rendimiento escolar, dado que estos constituyen una de las expresiones actuales más representativas de las dimensiones y evolución de aquellas.

Aunque desde sus orígenes la medición en educación ha estado marcada por una procura incesante de la posibilidad de comparar validamente sus resultados entre sujetos, escuelas y países, la que se ha plasmado en el desarrollo de la condición estandarizada de los mismos, el auge mundial de esta forma de medición es asunto de no más de veinte años. Sus expresiones previas, aun las más antiguas con casi un siglo de existencia, estuvieron restringidas casi exclusivamente a los Estados Unidos, surgiendo inicialmente en el campo de la psicología y trasladándose luego al de la educación.

Las primeras acciones para internacionalizar la evaluación estandarizada surgieron en la década del setenta con el primer estudio realizado por la Asociación Internacional para la Evaluación del Rendimiento Educativo (IEA)¹ siendo ésta entidad la que llevó a cabo la primera serie de estudios internacionales formales en la áreas de ciencia, literatura, comprensión lectora, Inglés y Francés como lenguas extranjeras y educación cívica, involucrando, en ese entonces, a diecinueve (19) países de todo el mundo (Comber y Keeves 1973).

FIGURA 1. NÚMERO DE PAÍSES PARTICIPANTES EN TRES ESTUDIOS INTERNACIONALES DEL RENDIMIENTO ESCOLAR, AGRUPADOS POR AÑOS



Como es posible apreciar en la Figura 1, la evolución en la participación de los países en algunos de los principales estudios internacionales de evaluación del rendimiento escolar ha crecido de manera apreciable desde la primera experiencia antes mencionada. En el caso del Programa Internacional de Rendimiento Estudiantil (PISA)², coordinado por la OCDE³, en un lapso de nueve años se observa un incremento en el número de países participantes cercano al sesenta por ciento. En la serie del programa

¹ International Association for the Evaluation of Educational Achievement.

² Programme for International Student Achievement.

³ Organización de Cooperación y Desarrollo Económico.

denominado Estudio de Tendencias en Matemática y Ciencias (TIMSS)⁴, desarrollado por la IEA y que se inició en 1970 con el denominado Primer Estudio de la IEA, antes mencionado, se aprecia un crecimiento en el número de países participantes de un ciento ochenta por ciento, en treinta y siete años. Finalmente, en el caso del Estudio Internacional de Progreso en Lectura (PIRLS), también a cargo de la IEA, el número de países participantes ha crecido en casi un veinte por ciento, en sólo cinco años.

Sin embargo como no siempre la demanda por un producto y su eficacia van juntas, cabe en este caso la pregunta: ¿puede el aumento en la participación nacional en estos estudios interpretarse como un aporte concreto al avance del aprendizaje de los alumnos -fin último de todo proceso que se concibe como educativo-? La respuesta a la pregunta recién planteada pasa antes por la de esta otra interrogante: ¿es la evaluación estandarizada en educación, en primer lugar, efectiva, -ello es cumple con los requisitos para lograr lo que se supone debe lograr- y, en segundo, de ser así, es ella eficaz, -lo logra en la realidad?

2. ¿POR QUÉ EFICACIA Y EFECTIVIDAD?

Dos de los principales requisitos que establecen el valor de cualquier acción son, en primer lugar, el que ella sea lo que dice ser, en otras palabras si es o no **efectiva** y si acaso como tal cumple con su propósito, en otros términos si es o no **eficaz**. De paso, es importante señalar que de hecho, la **efectividad** condiciona la **eficacia**.

Sin embargo, hay pocos términos en el campo de la educación cuyo sentido sea confundido con mayor frecuencia que el de estos dos. Se habla, por ejemplo, de “escuelas efectivas” cuando se quiere hacer referencia a “escuelas eficaces” y todo ello debido a una errónea traducción desde el idioma Inglés, del cual nos llegan la mayoría de las expresiones técnicas en educación. En efecto, tendemos a confundir los términos, ya que por una parte, la traducción correcta del término inglés “effectiveness” al castellano es eficacia -lo distintivo de aquello que cumple con su propósito- y por la otra, el término “efectivo” en castellano describe a aquello que es real.

3. EFECTIVIDAD Y EFICACIA VERSUS VALIDEZ Y CONFIABILIDAD

Son precisamente la efectividad y la eficacia los elementos que se ubican en la raíz misma de gran parte del debate acerca de la medición y la evaluación estandarizadas. Avanzando en este razonamiento si transferimos estos términos, provenientes del campo de la planificación al de la psicometría, nos encontramos con que la **efectividad** se relaciona mayoritariamente con la **validez**, condición que “grosso-modo” califica si un instrumento mide lo que pretende medir y **eficacia** se traduce primordialmente en **confiabilidad**, ello es define el grado en que tal instrumento mide adecuadamente. De forma equivalente a lo señalado anteriormente, en psicometría la **validez** condiciona la **confiabilidad**.

De los dos conceptos señalados, aquel al que este artículo pretende dedicar mayor atención es al de la validez. Existe hoy un debate acerca de si la definición originaria y más propia de esta cualidad de los instrumentos en educación, se enfocaba en un solo ámbito (validez de contenido) (Lissitz y Samuelsen, 2007) o si esa definición implicaba alcances más amplios (validez de constructo y validez concurrente)

⁴ Trends in Mathematics and Science Study.

(Moss, 2007). En términos concretos, si acaso al hablar de validez se aludía primariamente a la consistencia entre el instrumento y el mapa de los contenidos y/o competencias por medir que se hubiesen identificado como relevantes o si ella también incluía con igual importancia, entre otras, a la concordancia entre lo que esa herramienta mide y una estructura teórica pre-definida y/o con lo que miden otros instrumentos.

Es preciso señalar, primeramente, que más allá de las diferencias antes señaladas, no existen límites absolutos entre los diferentes ámbitos de la validez y que estos en la realidad se superponen. En segundo lugar, es importante dejar claro que la validez es una variable continua e incluso, las más de las veces imposible de cuantificar con precisión a diferencia de la confiabilidad y en tercero que no existen instrumentos totalmente válidos o inválidos, sino que estos pueden alcanzar diversos grados de validez no siempre nítidos. Finalmente, como lo señala una publicación conjunta de tres entidades profesionales norteamericanas (AERA⁵, APA⁶ y NCME⁷, 1989) citada por Moss (2007), es más propio hablar de **roles** de los variados tipos de evidencia de la validez o de formas de validación, que de **tipos** de validez propiamente tal.

Sin perjuicio de lo recién señalado y más allá de propugnar aquí la adhesión a una concepción unitaria y focalizada o desagregada y abierta de la validez, aunque éste autor se pronuncia más bien por la primera de ellas, se plantea a continuación un intento, de definición de los posibles tipos más básicos de evidencia en este sentido. Es así como en este aspecto es posible reconocer los siguientes ámbitos:

- La **validación por "constructo"**⁸ tiene lugar "...cuando un investigador sustenta que un instrumento refleja una particular estructura teórica o "constructo", a los cuales se atribuyen ciertos significados. La interpretación propuesta genera hipótesis verificables, las que constituyen medios de confirmar o rechazar el planteamiento (contenido en el "constructo")..." (Cronbach y Meehl, 1955, citado por Moss, 2007). Un ejemplo concreto corresponde al caso en que se pretende evaluar conocimiento de matemática con pruebas de lápiz y papel, en niveles de grado o curso donde la comprensión lectora está aún en estado de consolidación y por ello es válido sustentar que el "constructo" que mide la prueba no es el conocimiento de matemática sino que una mezcla de éste con la comprensión lectora.
- La **validación por contenidos** es aquella que se lleva a cabo verificando la medida en que una prueba "...constituye una muestra representativa de la clase de situaciones o materias acerca de las cuales se deben extraer conclusiones..." (American Psychological Association, 1966, citado por Moss, 2007). Ejemplo específico, también negativo y típico, es aquel en que los contenidos curriculares del instrumento no corresponden a la tabla de especificaciones y la prueba ya sea, mide lo que no está en la tabla o no mide lo que sí está.
- La **validación por concurrencia** verifica la medida en que hay coincidencia "cuando se comparan los resultados de una prueba con una o más variables externas que se considera que proveen una medida directa de la característica o conducta en cuestión" (American Psychological Association, 1954, citado por Moss, 2007).

⁵ American Educational Research Association (Asociación [Norte] Americana de Investigación Educativa).

⁶ American Psychological Association (Asociación [Norte] Americana de Psicología).

⁷ National Council on Measurement in Education (Consejo Nacional [de EEUU] para la Medición en Educación).

⁸ El término inglés "construct" no tiene una traducción precisa en castellano, siendo la que más se acercaría a su sentido conceptual la de "estructura teórica", por lo cual en este artículo se ha optado por utilizar el anglicismo "constructo" puesto entre comillas.

A estas definiciones primigenias, hay que agregar lo expresado en una versión de los estándares para la medición en psicología y educación de AERA, APA y NCME (1999), en la cual se menciona que, al fin de cuentas, "...existen cinco categorías de evidencia acerca de la validez, a ser consideradas al construir un argumento ..." "el contenido del instrumento, los procesos de respuesta, la estructura interna, las relaciones con otras variables y las consecuencias de la (aplicación de la) prueba". Es posible sustentar que éste planteamiento coincide con lo señalado por Messick (1989) acerca de que "La validez es un juicio integrado acerca del grado en el cual la evidencia empírica y las concepciones teóricas respaldan a las *inferencias* y *acciones* basadas en los resultados de las pruebas y otros modos de medición, en cuanto que (aquellas) sean *adecuadas* y *apropiadas* (cursivas en el texto original).

Estos últimos dos alcances tienen gran importancia ya que extienden significativamente la gama de evidencias que son necesarias para respaldar la validez, incluyendo aspectos externos tales como las consecuencias de la aplicación de los resultados de los instrumentos. Un reciente artículo de Nichols y Williams (2009) explora este tema en mayor profundidad aún, discutiendo temas tales como la atribución de responsabilidad a los distintos actores involucrados en la medición, por de los distintos tipos de evidencia relacionados con las pruebas en educación.

Una consecuencia práctica muy importante de destacar acerca de la validez, de acuerdo a lo indicado en el párrafo anterior, es que hoy en día se considera que ésta cruza una muy amplia gama de aspectos relacionados con la prueba a la que califica, los que van desde la fidelidad de la misma al mapa de contenidos o competencias que la originó, hasta los efectos que sus resultados pueden generar, pasando por combinaciones de ellos. Un buen ejemplo es el caso de las pruebas de base curricular, como las del estudio TIMSS que, aún cuando su estructura pueda mostrar una alta relación promedio con los currículos de la mayoría de los países, ésta suscitará cuestionamientos en términos de su validez para medir los rendimientos en países en los cuales no existe un currículo único, sino que la docencia se basa en estándares más ligados con competencias. En casos como éste la conveniencia de adherir a un estudio como PISA, basado en estándares de validez universal, aparecería como más plausible.

4. INTENTANDO DESCRIBIR EL "ICEBERG" DE LA MEDICIÓN EN EDUCACIÓN

Seguidamente, la exploración en mayor profundidad del tema de la validez de los instrumentos nos lleva a constatar una realidad que por más que pueda no ser del agrado de los especialistas en medición, no tiene mayor posibilidad de ser desmentida y más aún, es preferible que sea reconocida y explicada para contribuir a desvirtuar críticas en contra de la medición estandarizada en educación. Tales críticas son producto, por una parte, de que los especialistas, a veces, en nuestro entusiasmo por los avances psicométricos perdemos de vista las limitaciones de nuestro oficio y le atribuimos mayores capacidades de las que realmente tiene y por la otra, por cuanto muchos sujetos afectados por resultados adversos en estas mediciones tratan de desvirtuar tales resultados desacreditando las herramientas que los originaron.

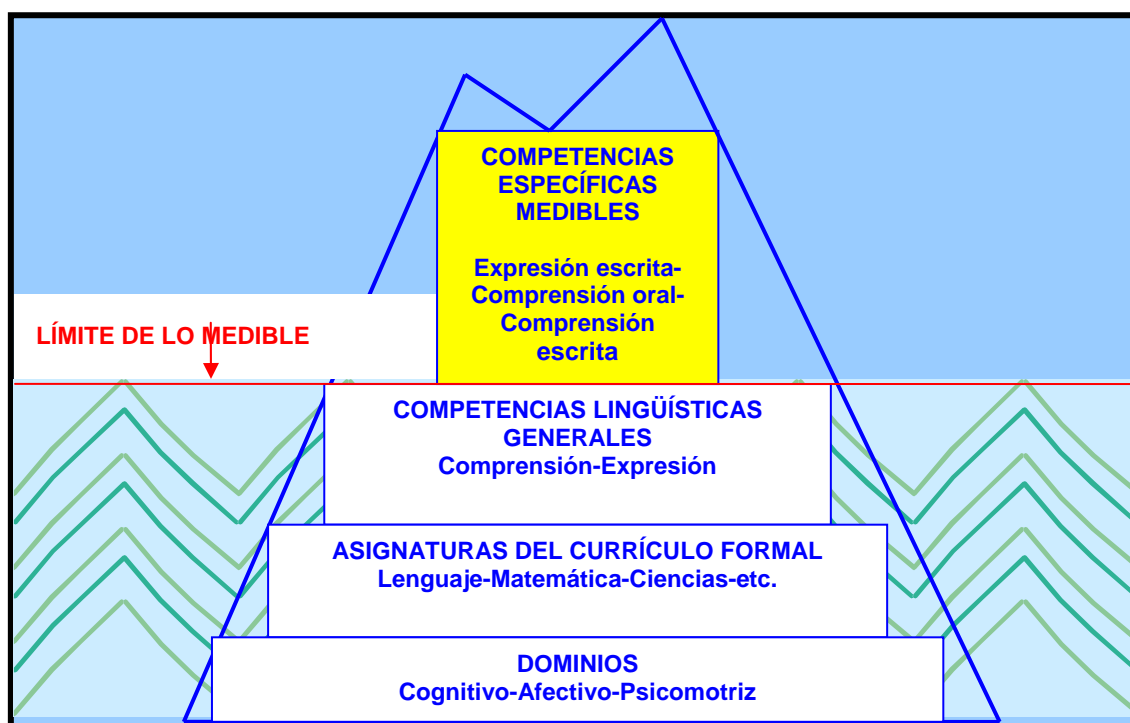
El caso es que las mediciones estandarizadas en educación están limitadas por diversos factores, siendo uno muy gravitante el que surge precisamente de su restringida validez relativa, producto de su naturaleza intrínseca de constituir un intento indirecto para detectar características internas de los sujetos a través de una muestra de evidencias externas. Es importante señalar que si bien este efecto es posible de atenuarse no logra, como podría pensarse, ser totalmente corregido por medio de un uso más intenso

de la tecnología, como podría ser el caso del CA⁹ y del CAT¹⁰, dado que su limitación es fundamental y reside, como se dijo antes, en que es un mecanismo indirecto y externo para detectar rasgos internos. Sin embargo, pese a sus limitaciones, hasta hoy no existen métodos científicamente aceptables que hagan un mejor trabajo, para estos efectos, que la medición estandarizada.

Es por esta razón que la proporción del total del aprendizaje, que realmente es posible medir con las evaluaciones estandarizadas corresponde sólo a la cima del "iceberg", que es aquella parte que se ubica por sobre el límite que separa a las características mensurables de forma estandarizada de las que no lo son, que al igual que en el caso de los témpanos que flotan en los mares fríos, corresponde sólo a una parte menor de su volumen total. Esto equivale, en el caso de la educación, también a una expresión minoritaria de lo que realmente ella debió haber logrado en los estudiantes.

Un análisis descriptivo relativamente simple permitirá apreciar esto más claramente si se toma como ejemplo la medición estandarizada de los aprendizajes en Lenguaje. Si se considera la educación formal, ésta abarca al menos tres dominios de competencias, mostrados en el nivel de la base de la Figura 2, por lo cual todo aprendizaje, realmente educativo, debe incluir a estos tres dominios aunque en proporciones variables. Es por ello que al utilizar la medición estandarizada se está capturando información de los dominios (cuando es posible hacerlo) de manera artificialmente segregada. Esto implica que en el ejemplo del Lenguaje, la medición estandarizada habitualmente sólo captura los aprendizajes cognitivos.

FIGURA2. EL "ICEBERG" DE LA MEDICIÓN EN EDUCACIÓN. AQUELLO QUE ES REALMENTE POSIBLE DE SER MEDIDO EN LENGUAJE, POR MEDIO DE PRUEBAS ESTANDARIZADAS ORDINARIAS



⁹ CA. Computerized testing (prueba computarizada). Situación en que la prueba es administrada a través de un computador, la que permite el uso de una mayor variedad de estímulos y por ello la medición de competencias más complejas, fundamentalmente usando técnicas de simulación.

¹⁰ CAT. Computerized adaptive testing (prueba adaptativa computarizada). Situación que incluye todas las ventajas del CA, pero que además posibilita adaptar las preguntas subsiguientes en base de las respuestas a las previas de cada sujeto que responde, logrando así una mayor coincidencia entre la prueba y las habilidades del examinado y con ello una medición más precisa.

Ascendiendo un nivel desde la base se aprecia que dentro del espectro cognitivo el aprendizaje involucra a varias asignaturas las cuales, si bien como todo aprendizaje humano también interactúan, son las más de las veces abordadas de manera también separada, por lo cual el Lenguaje como asignatura representa también sólo una fracción, esta vez del aprendizaje cognitivo formal total.

Yendo a un nivel adicional más arriba y entrando en el área curricular del Lenguaje se aprecia que existen al menos dos tipos de competencias generales que debieran ser parte de cualquier diseño curricular en la asignatura: la comprensión y la expresión. Y ambos debieran medirse si la intención es verificar los aprendizajes en esta área curricular.

Sin embargo, al abordar estas dos competencias básicas es posible comprobar que las competencias específicas posibles de medir de forma estandarizada son solamente las siguientes: la expresión escrita, a través de preguntas de respuesta abierta; la comprensión oral, por medio de ítems cerrados o abiertos; y la comprensión escrita, por estos mismos medios. Aunque algunos autores (Rodríguez, 2002) hacen depender la elección del tipo de ítem de aspectos tales como las recomendaciones de la comunidad de especialistas, el costo e incluso de las repercusiones del uso de un tipo u otro en términos de la política educacional, ella siempre dependerá antes que nada del nivel de complejidad de las competencias que se busque medir. Las más de las veces se opta, debido a razones de costo, por escoger sólo ítems cerrados, con la consiguiente reducción en la validez. Similarmente, la expresión oral se deja fuera debido a la complejidad y costo, si es que se quiere garantizar su objetividad, de medirla en forma estandarizada.

En conclusión, de tres dominios se mide uno, sucediendo lo mismo con el espectro de asignaturas incluidas y dentro de la única abordada, de sus dos competencias generales se logra medir parcialmente una y de forma relativamente integral la otra.

Si bien es posible complementar la medición estandarizada con otros enfoques como la observación y los portafolios, estos presentan limitaciones a la hora de tratar de comparar sus resultados entre alumnos de distintas escuelas y aulas, ya que no constituyen, de suyo, métodos estandarizados y si bien es posible aproximarlos a tal condición, ello implica, una vez más, un aumento sustancial del tiempo y del costo involucrados.

5. MEDICIÓN Y EVALUACIÓN ¿DISTINTOS TÉRMINOS PARA UN SOLO CONCEPTO?

Para contribuir a develar la complejidad del tema abordado, otro aspecto que es importante precisar es que dos términos que en su ámbito se usan regularmente de manera indistinta no son, en rigor, sinónimos. Se trata de **medición** y **evaluación**.

Si bien existen múltiples definiciones de ambos, una forma directa de percibir sus diferencias es considerando que la medición consiste en el acto de establecer la posición de un sujeto en la escala de una variable determinada y por lo tanto constituye una visión instantánea y, por ello, estática, de su situación respecto de esa variable. Evaluación, por su parte, implica determinar la trayectoria que, en el tiempo, muestran las posiciones de un sujeto en la escala de una misma variable, por lo cual implica una visión dinámica del comportamiento que tal sujeto ha tenido en esa variable, obtenida por medio de la comparación de los resultados de dos o más mediciones de ese mismo sujeto.

Por lo recién señalado, de acuerdo a esta visión de ambos procesos, la medición constituye un insumo de la evaluación y su calidad determina en gran medida la de ésta. En este último sentido aunque los

métodos de comparación de las varias mediciones sean óptimos, si la precisión, objetividad, validez y confiabilidad de éstas no satisfacen estándares aceptables, el valor de la evaluación será cuestionable.

6. HERRAMIENTAS PARA DIFERENTES FINES: REFERENCIA A NORMAS O A CRITERIOS

Los primeros instrumentos de medición estandarizada surgieron de la necesidad de clasificar u ordenar sujetos entre sí y se denominan pruebas **referidas a normas**. Una forma simple de visualizar el mecanismo a través del cual éstas operan es considerar que lo que hacen es estructurar la distribución de los resultados obtenidos por los sujetos en un instrumento, de forma relativa y luego establecer el ordenamiento de estos últimos de manera acorde. Es importante precisar que en este enfoque la distribución de los resultados necesariamente mostrará proporciones de sujetos en las distintas categorías de resultados coincidentes con una distribución típica, como es la normal, ello es que porcentajes menores de los resultados corresponderán a sujetos con muy altos o muy bajos puntajes y que la mayor proporción de los resultados de los examinados se ubicarán en las áreas de puntajes intermedios. Un ejemplo de este tipo de instrumentos son las pruebas de selección para el ingreso a la educación superior, conocidas, en los Estados Unidos, como Scholastic Aptitude Test (SAT)¹¹.

El segundo enfoque, de desarrollo posterior al anterior, incluye instrumentos cuyo propósito es determinar la posición de los sujetos respecto de un referente externo, absoluto y pre-establecido, no considerando especialmente relevante la posición relativa de los mismos. En esencia lo que interesa establecer, en este caso, es cuáles sujetos pasan una determinada valla, criterio o puntaje, para establecer si son o no competentes de manera absoluta en un aspecto determinado. Este tipo de instrumentos se denominan **referidos a criterios** y en su desarrollo e interpretación ha contribuido significativamente la denominada Teoría de Respuesta al Ítem (TRI)¹². Las pruebas de rendimiento corresponden a este tipo de instrumentos y un buen ejemplo lo constituyen las usadas por los estudios internacionales mencionados al comienzo de este artículo. Es importante señalar que la distribución de los resultados de este tipo de pruebas no tiene por qué corresponder a la normal y más aún, es esperable que no lo haga, ya que si se considera que la educación es un proceso de crecimiento equitativo para todos los estudiantes, sería esperable que su aprendizaje no se distribuyera normalmente.

7. ¿PUEDEN LOS MISMOS INSTRUMENTOS CUMPLIR UNA DIVERSIDAD DE ROLES?

La necesidad que existe, en muchos países, de reducir los costos de procesos de suyo onerosos y de grandes dimensiones, como son los de medición y evaluación, como asimismo la conveniencia de disminuir la frecuencia con que los estudiantes son sometidos a mediciones -baste considerar que al menos tienen que enfrentar las del docente, las del sistema nacional y las internacionales- ha llevado a intentar en algunos casos la utilización de los mismos instrumentos para ordenar a los sujetos y para determinar su nivel de aprendizaje. Un ejemplo fue el Sistema de Ingreso a la Educación Superior (SIES), en Chile, diferido en su aplicación desde hace algunos años.

¹¹ En castellano Prueba de Aptitud Académica.

¹² La sigla TRI en castellano deriva de la inglesa "IRT", la que corresponde a Item Response Theory.

Como señalara Kaplan (1963) "el puntaje de una prueba puede entregar interpretaciones relativas o absolutas." Por su parte, abordando el tema desde una perspectiva sutilmente diferente de la utilizada en la sección anterior, Haladyna (2002), agrega que "la interpretación absoluta es denominada usualmente *referida a criterios* (RC) debido a que nos indica cuánto de una habilidad o conocimiento ha adquirido una persona". "La interpretación relativa es llamada *referida a normas* (RN), por cuanto esta interpretación permite la comparación de un puntaje con otros puntajes en la escala de una prueba determinada. Las interpretaciones RN nos indican cuan diferentes unos de otros son los puntajes."

Citando nuevamente a Haladyna (2002), es posible señalar "que aunque RC y RN corresponden a tipos de interpretaciones de los puntajes de una prueba, en sentido estricto no tenemos pruebas RC y RN. Sin embargo, los términos RC y RN se asocian a menudo con pruebas que permiten tales interpretaciones." Y agrega que "Irónicamente cualquier prueba puede aportar interpretaciones NR, pero las interpretaciones CR provienen (sólo) de pruebas diseñadas para informarnos cuánto ha aprendido un estudiante."

Si bien ésta es, como se dijo antes, una forma leve pero fundamentalmente distinta para la conceptualización de las diferencias entre los enfoques referidos a normas y a criterios, en esencia la conclusión es la misma en cuanto a que ambos tipos tienen propósitos distintos. El punto siguiente abordar es la aclaración de si acaso tal diferencia descalifica a un tipo para cumplir el rol del otro, o si es posible pensar en una figura híbrida.

Por una parte, la afirmación antes citada acerca de que sólo las pruebas CR pueden aportar información acerca del aprendizaje real de los estudiantes ya indica que, por definición y construcción, éstas son capaces de aportar elementos que las NR no son, por lo cual éste ya constituye un elemento de juicio a considerar en cuanto a que las pruebas NR no están en condiciones de aportar a la medición de los aprendizajes respecto de criterios absolutos.

Por otra parte, cada uno de estos tipos de interpretaciones enfatiza ciertas características que hacen que en la construcción de las pruebas mismas se privilegien ciertos aspectos en desmedro de otros. Tal es el caso de que las pruebas CR tienden a poner el acento en la validez y las NR en la discriminación.

El caso es que, además, los dos tipos de herramientas descritas en la sección anterior se basan en modelos de interpretación diferentes, descansan en supuestos estadísticos distintos (como es el de la normalidad de la distribución de los puntajes en el caso de las NR y no necesariamente en el de las CR) y finalmente entregan información diferente.

En conclusión y en una comparación que puede ser considerada trivial, la utilización indistinta de cualquier de los tipos para cumplir ambos propósitos -jerarquización y verificación del aprendizajes de los sujetos, medidos- puede tener el mismo efecto que emplear un coche de ciudad a campo traviesa, caso en el cual es posible que éste logre efectuar el recorrido, pero con serios riesgo para su integridad y la de sus tripulantes.

8. FORMAS DE EXPRESAR PUNTAJES Y FACILITAR SU COMPARABILIDAD

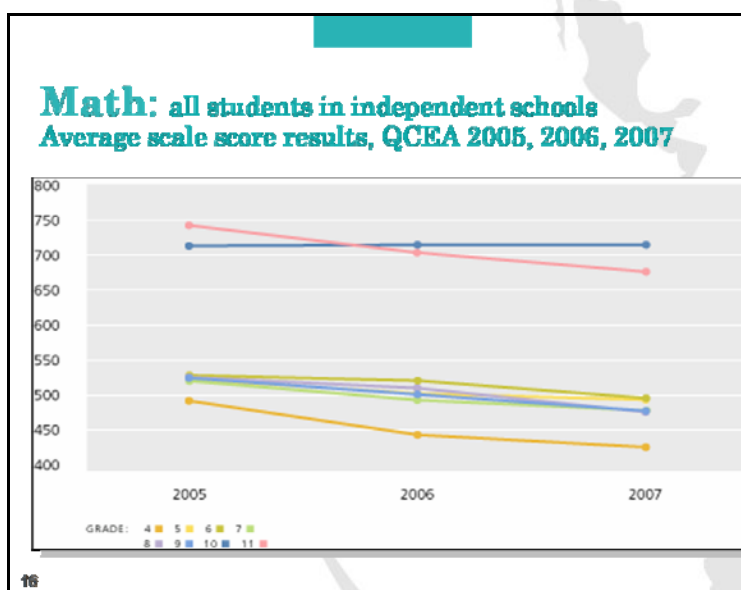
El aspecto que se busca explorar en esta sección es la facilidad relativa que implica la comparación de resultados expresados, ya sea de forma cuantitativa o cualitativa, sin entrar aún al tema de la comparabilidad estadística de los mismos.

Como se ha mencionado repetidamente en este artículo uno de los aspectos que más ha preocupado y sigue preocupando hoy en día, tanto a los especialistas en medición como a los educadores en general, es la posibilidad de comparar los resultados de las mediciones estandarizadas, posibilitando de esta forma una verdadera evaluación de sujetos, aulas, escuelas o países.

Estas comparaciones, las más de las veces basadas en el cálculo de diferencias entre los resultados de dos o más mediciones, pueden ubicarse a distintos niveles de agregación. La forma de comparación puede ser longitudinal, implicando que los resultados de los mismos sujetos son contrastados en distintos momentos. En otros casos las comparaciones pueden ser transversales, lo que significa que se busca comparar a distintos sujetos, en los mismos niveles de grado, en distintos momentos. Existe también la posibilidad de comparar sujetos en distintos niveles de grado, aunque ello requiere del cumplimiento de determinadas condiciones psicométricas. Finalmente, el método del Valor Agregado, consistente en la comparación de un valor de rendimiento predicho y uno efectivo, puede también aplicarse en estas contrastaciones.

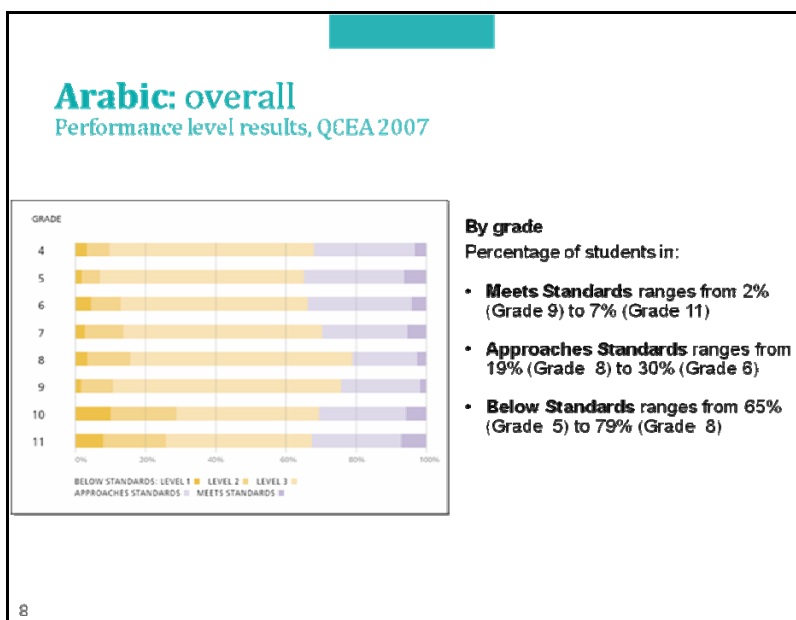
La forma cuantitativa más ortodoxa de presentación de puntajes es en escala IRT para las pruebas referidas a criterios y para estas mismas pruebas la forma aceptada de presentar resultados en forma cualitativa es por medio de niveles de desempeño, lo cual implica la previa determinación de estos por medio de cualesquiera de los métodos universalmente aceptados, tales como el Angoff, el de Marcadores, el de Juicio Analítico u otros (Cizek, 2001).

FIGURA 3. COMPARACIÓN LONGITUDINAL DE RESULTADOS ENTRE LOS AÑOS 2005, 2006 Y 2007, EXPRESADOS EN ESCALA CUANTITATIVA, EN MATEMÁTICA, POR NIVELES DE GRADO, COMO PARTE DEL SISTEMA DE MEDICIÓN EDUCACIONAL DEL ESTADO DE QATAR (QCEA). (SEC, 2007)



La posibilidad de comparación de resultados numéricos, sea ésta vertical u horizontal, está condicionada por la equivalencia psicométrica de los instrumentos. En cambio la comparación de resultados expresados en clave cualitativa es mucho más flexible. A continuación se muestran dos ejemplos: uno correspondiente a resultados expresados de forma cuantitativa (Figura 3) y el otro en forma cualitativa (Figura 4). Ambos ejemplos corresponden al programa denominado Medición Integral en Educación, en el Estado de Qatar¹³, en la Península Árabe.

FIGURA 4. COMPARACIÓN DE RESULTADOS ENTRE GRADOS PARA EL AÑO 2007, EXPRESADOS EN ESCALA CUALITATIVA (NIVELES DE DESEMPEÑO), EN LENGUAJE (ÁRABE), COMO PARTE DEL SISTEMA DE MEDICIÓN EDUCACIONAL DEL ESTADO DE QATAR (QCEA)



Como es posible apreciar, las escalas cualitativas permiten una mayor flexibilidad de comparaciones y una mucha mayor riqueza de información, sin requerir el cumplimiento de condiciones tan estrictas como en el caso de las escalas cuantitativas, aunque lógicamente no logran una precisión comparable con el nivel posible de alcanzar al usar escalas cuantitativas. Sin embargo, incluso es posible, aunque no necesariamente aconsejable, utilizar la información acerca de porcentajes de alumnos en cada nivel para comparar el rendimiento entre distintas asignaturas, lo que constituye un anatema siquiera intentarlo a partir de resultados cuantitativos.

9. DESDE LAS CIFRAS A LOS CONCEPTOS

Como se desprende de la sección anterior, existen razones y muy poderosas, para que se justifique un mayor desarrollo y énfasis en el uso de resultados en escala cualitativa, a la hora de diseminar y analizar los resultados de las mediciones estandarizadas en educación.

¹³ Qatar Comprehensive Educational Assessment (QCEA) program

Sin embargo, es necesario señalar claramente que ello no implica licencia para concesión alguna en el rigor de la generación de resultados o para la prescindencia de las escalas cuantitativas, sino que significa tomar ventaja de la mayor flexibilidad de posibilidades de comparación planteada por los resultados expresados en forma de niveles de desempeño, generados por medio de la determinación de puntajes de corte, a partir de resultados cuantitativos, idealmente en escala IRT.

Los procesos de determinación de niveles de desempeño están, hoy en día, definidos y estandarizados en tal medida, que es preciso seguir pasos muy claros para el establecimiento de los niveles y posteriormente de sus límites, a través de la determinación de puntajes de corte. Aunque estos procesos descansan en procedimientos de juicio y no necesariamente en unos de raíz estadística, sí implican un rigor tanto o mayor que el propio de estos últimos.

La posibilidad de contar con información cualitativa acerca de los resultados de las mediciones estandarizadas en educación permite alcanzar tres fines. El primero, como se señaló antes, consiste en la ampliación del ámbito de los tipos de comparaciones posibles de realizar a partir de los resultados de las mediciones estandarizadas. El segundo es la asignación, a tales comparaciones, de un grado mayor de inteligibilidad para los destinatarios de la información, tales como directores, maestros, padres y alumnos, despejando la habitual duda acerca de cuál es el sentido que, en términos educativos, tiene un puntaje numérico determinado. Finalmente, establece un puente entre los resultados de las mediciones y la docencia de aula, en el sentido de que al establecer cuáles son las competencias concretas que los alumnos logran y cuáles las que no alcanzan, se genera una herramienta efectiva de retro-información y por ende de mejoramiento del aprendizaje.

La integración, actualmente posible, de resultados cuantitativos, generados con todo el rigor psicométrico propio de ellos, con esquemas de interpretación cualitativos generados también a partir de procesos igualmente rigurosos, plantea una línea de explotación de los resultados de las mediciones estandarizadas, ausente hasta ahora en el ámbito educacional, pero que promete nuevas y grandes posibilidades en beneficio de los estudiantes.

10. PREDICCIÓN, ALINEAMIENTO Y "ECUALIZACIÓN" DE PUNTAJES

Como se adelantara dos secciones más atrás en este artículo, la respuesta a la necesidad de comparar resultados generados por las mediciones estandarizadas en educación, ha constituido y constituye actualmente, una necesidad que hace posible el cumplimiento de un propósito fundamental de la medición, cual es la evaluación.

La principal razón de esta necesidad surge del hecho que, las más de las veces, lo que se intenta contrastar son distintos elementos (sujetos) en un mismo o distinto momento (comparaciones transversales) o los mismos elementos (sujetos) en diferentes momentos (comparaciones longitudinales). En ambos casos surge como condicionante la necesidad de que los instrumentos usados sean comparables y para ello se utilizan procedimientos psicométricos denominados de "vinculación"¹⁴.

¹⁴ En inglés "linkage".

No son pocos los casos en que erróneamente se han generado conclusiones alentadoras o negativas respecto de la situación de sistemas escolares, producto de que se ha intentado comparar puntajes que no son comparables, dado que estos provenían de instrumentos que no eran equivalentes.

La denominada “vinculación” de pruebas hace su aparición a comienzos del siglo veinte, con los trabajos de varios autores, entre ellos, Starch, Weiss y Kelley, citados por Holland (2005).

La vinculación de puntajes de pruebas, de acuerdo al mismo autor recién citado, comprende tres grandes sub-categorías, a saber:

- **Predicción.** Lo que se obtiene es la “mejor predicción”¹⁵ de un puntaje Y a partir de un puntaje X, utilizando procedimientos de regresión.
- **Alineación de escalas**¹⁶. Es un proceso más complejo que la predicción y consiste en la ubicación conjunta de los puntajes X e Y en una escala común, en vez de asignar a X un puntaje en la escala de Y. Los métodos más comunes son la alineación de batería, la alineación respecto de un ancla¹⁷, la alineación respecto de una población hipotética, la calibración, la alineación vertical y la concordancia.
- **“Ecuación”**¹⁸. Es el proceso más complejo de todos y se lleva a cabo cuando los diferentes sujetos rinden diferentes formas o versiones de la misma prueba y sus resultados en ellas requieren ser comparados. Este proceso es posible abordarlo también con Teoría Clásica, IRT, ecuación por post-estratificación, ecuación en cadena, método de puntajes observados de Levine y método IRT de puntajes observados.

Para finalizar esta breve revisión acerca del origen y de la práctica de los métodos que permiten la comparación válida de puntajes de instrumentos de medición del rendimiento, es preciso mencionar que una de las modalidades comprendidas dentro de los métodos de alineamiento, debe destacarse un método que recientemente ha ganado en popularidad, cual es la alineación vertical.

Su mayor mérito radica en que hace posible la estimación del crecimiento o progreso del aprendizaje de los alumnos en el tiempo, entregando información, por ejemplo, de cuál ha sido la trayectoria de un alumno en matemática desde tercero a cuarto de educación básica. Un método alternativo y menos exigente en términos psicométricos es el cálculo de valor agregado, aunque sí lleva aparejadas otras exigencias, esta vez operacionales. En ambos casos la limitación mayor es la “distancia” en el tiempo que pueda existir entre ambas mediciones, la que se expresa en el número de niveles de grado en el caso de la alineación vertical y en el tiempo transcurrido entre la mediciones en el del valor agregado.

11. CONDICIONES PARA QUE UNA EVALUACIÓN CONSTITUYA UNA HERRAMIENTA DE POLÍTICA EN EDUCACIÓN

Uno de los aspectos más controversiales de la medición en educación es su utilización para la toma de decisiones en materias de política educacional, dado que el volumen de datos y de información que se acumula en los anaqueles de los ministerios y otras entidades de los sistemas educacionales es enorme,

¹⁵ En inglés “best prediction”.

¹⁶ En inglés “scale alignment” o “scaling”.

¹⁷ En inglés “anchor test”.

¹⁸ En inglés “equating”.

como asimismo lo son los bancos de datos que rebosan de resultados de operaciones de medición. Es también muy grande la carga que se impone sobre profesores y alumnos a raíz de someter a estos últimos a evaluaciones, generadas al nivel escuela, municipio, provincia o estado y/ país y no devolverles mayores beneficios.

Sin embargo, el número de ocasiones en que las decisiones en materia de política educacional son basadas en la información dura, generada por la medición estandarizada, es inversamente proporcional al crecimiento del volumen de información generada. Si bien autores como Gladwell (2005) abogan por el valor de las decisiones producto de la inspiración súbita, cabe la duda de si acaso aquellas en educación puedan y sea aconsejable que sean adoptadas de tal modo.

Lamentablemente las razones por las que la información generada a partir de los procesos de medición estandarizada no es utilizada en favor de decisiones mejor fundadas en educación es, en gran medida, responsabilidad de quienes estamos involucrados en tales procesos. Un análisis, a la vez sistemático y descarnado, de la situación imperante en este sentido, nos permitirá aislar las causas principales y por ese medio intentar subsanarlas.

La primera condición, que no siempre se cumple, es la que constituye la raíz del tema de este trabajo, la medida en que la información generada por la medición estandarizada es válida. En otras palabras, mientras las pruebas sólo se enfoquen en las competencias más simples de la escala de complejidad de los aprendizajes, su relevancia será discutible y por ende el estímulo implícito para la utilización de sus resultados será también limitado. Es imprescindible que los sistemas de medición desplieguen esfuerzos por ampliar el ámbito de validez de sus instrumentos, aún en tiempos en que los presupuestos no son del todo abundantes.

Una segunda condición se relaciona con la necesidad de que la información proveniente de las mediciones estandarizadas esté disponible para los diferentes actores de manera oportuna y sea inteligible para ellos. Son muchos los casos en que los informes de las mediciones toman años en ver la luz, cuando la ven, y lo que es más, suelen estar escritos en código de especialistas, siendo por ello incomprensibles para docentes, padres y alumnos y consecuentemente no concitan su interés.

Las más de las veces se invierte un enorme esfuerzo en la elaboración y análisis de los instrumentos de una medición, lo cual es sin duda lo pertinente de hacer, sin embargo no se asigna a la elaboración y producción de los informes el mismo esfuerzo e interés. Las Figuras 5 y 6 muestran un ejemplo de un informe para profesores que, desde hace dos años, está siendo provisto a los docentes en el Estado de Qatar, basándose en los resultados del sistema de medición del rendimiento, en ese país. Tal informe, junto con el correspondiente para padres y alumnos, el informe para los tomadores de decisiones a nivel superior y el informe técnico para especialistas, se procura que esté en sus manos no después de seis meses de aplicadas las pruebas, lo cual no si bien no es un plazo breve, sí implica una voluntad de hacer llegar la información en la forma lo más oportuna y útil posible.

FIGURA 5. PORTADA Y CONTRAPORTADA DEL INFORME PARA DOCENTES DE LENGUAJE (ÁRABE), UTILIZADO POR EL SISTEMA DE MEDICIÓN DEL RENDIMIENTO DEL ESTADO DE QATAR, PARA EL PERÍODO 2005-2006. (SEC QATAR, 2007)

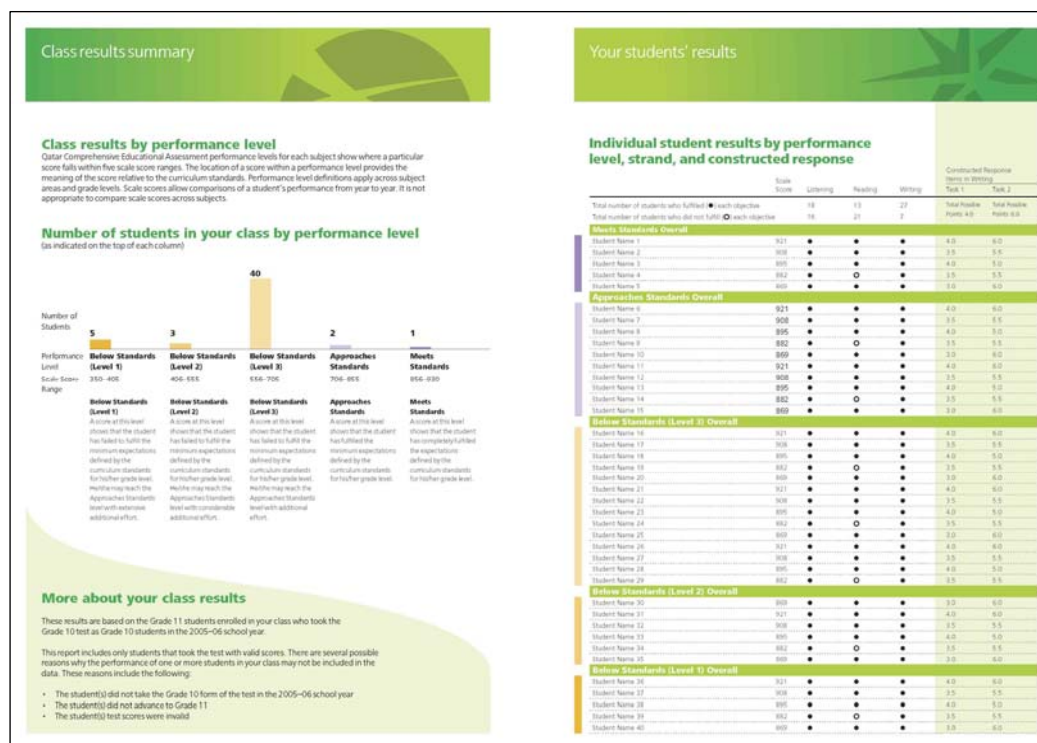


Otra condición limitante en la utilización de los resultados de las mediciones estandarizadas es la falta de desarrollo de la capacidad y voluntad para utilizar la información de quienes toman las decisiones en educación. De hecho, quienes deciden en este ámbito, desde padres hasta ministros, no están, las más de las veces, capacitados para utilizar información dura para sustentar sus juicios. Contra esta capacidad conspira también lo que se manifestara antes, en esta misma sección, en cuanto a que no siempre el lenguaje de los informes resulta adecuado para los actores en cada nivel que lo recibe.

Pero habiendo subsanado lo anterior deben hacerse esfuerzos para generar acciones concretas en pro de que estos actores, por una parte reciban capacitación para utilizar los informes basados en la medición estandarizada como base para sus decisiones, como asimismo para motivarlos a que lo hagan.

Lo más habitual hoy en día es que los padres tomen decisiones respecto de la educación de sus hijos basados en rumores y comentarios de sus conocidos. En el caso de quienes deciden en los niveles altos de los sistemas el caso es que las presiones de grupos relacionados con la educación tienen más efecto en lo que se resuelve que la información proveniente de fuentes objetivas y empíricas. Mientras esto siga ocurriendo, como lo señalara Bloom, en la cita al comienzo de este trabajo, seguiremos "...siendo seducidos por el equivalente de los remedios de curanderos, las falsas curas del cáncer, la invención del movimiento perpetuo y las supersticiones...".

FIGURA 6. PÁGINAS INTERIORES DEL INFORME PARA DOCENTES DE LENGUAJE (ÁRABE,) UTILIZADO POR EL SISTEMA DE MEDICIÓN DEL RENDIMIENTO DEL ESTADO DE QATAR, PARA EL PERÍODO 2005-2006. (SEC QATAR, 2007)



12. MITOS Y ELABORACIONES RETÓRICAS: ¿LA "APARICIÓN" DE LA EVALUACIÓN FORMATIVA? ¿LA MEDICIÓN DEL APRENDIZAJE VERSUS LA MEDICIÓN PARA EL APRENDIZAJE?

Uno de los aspectos que se mencionó antes como problemático respecto de la medición estandarizada de los resultados de la educación, en la actualidad, es la mítica aparición de "nuevas" tendencias y enfoques en este campo. Si bien los ejemplos son numerosos, un caso emblemático es el de la evaluación formativa. Es sorprendente como para muchos "especialistas" en evaluación y actores en educación, hoy en día, éste es un término que llegó muy recientemente a este campo por primera vez. O bien no estuvieron presentes cuando estos términos vieron la luz por primera vez, o estaban distraídos en ese entonces o simplemente no han hecho sus deberes para la casa.

Lo concreto es que, como señala una autora (Dwyer, 2008) "En 1967, Michael Scriven propuso el uso de los términos "formativo" y "sumativo" (comillas en el original) para distinguir entre los diferentes roles que puede jugar la evaluación. Por una parte, él señaló que la evaluación "puede tener un rol en el mejoramiento continuo del currículo" (p.41), mientras que por la otra, la evaluación "puede servir para habilitar a los administradores para decidir si el currículo completo y terminado, refinado por el uso del proceso evaluativo en el primero de sus roles, representa un avance suficientemente significativo entre las alternativas disponibles, para justificar su adopción por un sistema escolar (determinado) (pp. 41-42). Luego propuso "utilizar los términos evaluación "formativa" y "sumativa" para calificar a los roles de la evaluación".

"Dos años más tarde, Benjamin Bloom (1969) aplicó la misma distinción a las pruebas usadas en el aula."

Como quedó demostrado, la evaluación formativa fue definida por Scriven, desde una perspectiva sistémica, como la herramienta para el mejoramiento continuo del currículo y, por Bloom, desde un punto de vista pedagógico del aula, como un elemento vital para mejorar el aprendizaje de los alumnos, en ambos casos hace más de cuatro décadas.

Hoy se intenta utilizar este concepto para desacreditar a la medición estandarizada, como asimismo a la evaluación que se basa en ella, ya que se las considera como opuestas a lo "formativo". Baste decir que, por un lado, la idea de lo formativo califica las acciones evaluativas, por lo cual ellas no son más o menos formativas en su esencia, sino que en su uso. Sería como decir que un bisturí es esencialmente malo por cuanto puede usarse para matar, además de ser una herramienta curativa. Por otro lado, uno de los artífices más distinguidos de la introducción del rigor en la medición en el aula fue, precisamente, Benjamin Bloom. Mal pudo él propugnar que la medición estandarizada pudiese percibirse como antagónica del rol formativo de la evaluación.

Entrando, ahora, al ámbito de la manipulación retórica de los términos en este campo, segundo fenómeno de reciente aparición, el caso más destacado es el de la distinción artificial y artificiosa entre evaluación **del** aprendizaje y evaluación **para** el aprendizaje, las que se hacen aparecer como antagónicas.

Un simple análisis lógico baste para desvirtuar este intento sofista. En primer lugar, cualquier procedimiento y entre ellos, uno evaluativo, antes que nada debe poseer un objeto sobre el que se lleva a cabo, el que en este caso es el aprendizaje. Por lo tanto, no es posible que exista una acción evaluativa "in abstracto", sino que esta debe evaluar algo en concreto.

En segundo lugar, esa misma acción debe tener un propósito, en este caso la mejora del aprendizaje posterior y éste propósito no tiene razón para ser considerado opuesto al objeto antes definido. Por lo señalado, es perfectamente aceptable que la evaluación del aprendizaje en un cierto momento, vaya en pro del aprendizaje posterior.

A lo señalado, se agrega que una revisión mínima de la historia de la medición estandarizada y de la evaluación ligada a ella, permitirá comprobar que, desde al menos los años sesenta del Siglo XX, en que surge en la educación la idea de retro-información, la evaluación ha sido definida como una herramienta de mejoramiento del aprendizaje. Cualquier otra elaboración retórica no es más que eso, ya que revela desconocimiento del campo de la disciplina o bien una falacia en el razonamiento, que va a parejas con lo negativo de la intención.

Al finalizar este trabajo, sólo resta señalar que se ha procurado mostrar aquí que, si bien la medición estandarizada y su consecuente, la evaluación, distan mucho de ser exhaustivas e irreprochables en su validez y precisión, no existe, hoy por hoy, un enfoque científicamente sustentado que permita plantear una alternativa válida.

La conclusión última es entonces que lo sabio y prudente de hacer es trabajar para mejorar estas dos herramientas de la educación y ser, a la vez, extremadamente cuidadosos en que al aplicarlas se evite, a toda costa, asignarles mayores cualidades que las que realista y científicamente demuestren poseer.

REFERENCIAS BIBLIOGRÁFICAS

- American Educational Research Association, American Psychological Association y National Council on Measurement in Education. (1985). *Standards for educational and psychological testing*. Washington: American Educational Research Association.
- American Psychological Association (1966). *Standards for educational and psychological tests and manuals*. Washington: APA.
- American Psychological Association (1954). Technical recommendations for psychological tests and diagnostic techniques. *Psychological Bulletin [Supplement]*, 51 (2, part 2).
- Bloom, B.S. (1972). Innocence in Education. *School Review*, 80, pp.333-352.
- Bloom, B. S. (1969). Some theoretical issues relating to educational evaluation. En R.W. Tyler (Ed.), *Educational Evaluation: new roles, new means. The 63rd yearbook of the National Society for the Study of Education*, 69, pp. 26-50.
- Cizek, J.G. (2001). *Setting performance standards. Concepts, methods, and perspectives*. Mahwah Erlbaum.
- Comber, L.C. y Keeves J.P. (1973). *Science education in nineteen countries*. Stockholm: Almqvist & Wiksell.
- Cronbach, L.J. y Meehl, P.E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, 52, pp. 281-302.
- Dwyer, C.A. (2008). *The future of assessment. Shaping teaching and learning*. New York: Erlbaum.
- Gladwell, M. (2005). *Blink. The power of thinking without thinking*. London: Penguin.
- Haladyna, T.H. (2002). Supporting documentation: assuring more valid test interpretations and uses. En Tindal y T.M. Haladyna (Eds.), *Large scale assessment programs for all students*. Mahwah: Erlbaum.
- Holland, P.W. (2005). Linking scores and scales: some history and organizing ideas. En *Linking and aligning scores and scales. A conference in honor of Ledyard R. Tuckers's approach to theory and practice*. Princeton: Educational Testing Service.
- Kaplan, A. (1963). *The conduct of inquiry: Methodology for behavioral science*. New York: Harper & Row.
- Lissitz, R.W y Samuelsen K. (2007). A suggested change in terminology and emphasis regarding validity an education. *Educational Researcher*, 36, pp. 437-448.
- Messick, S. (1989). Validity. En R.L. Linn (Ed.), *Educational Measurement*. New York: Macmillan.
- Moss, P.A. Reconstructing validity. (2007). *Educational Researcher*, (36, pp. 470-476)
- Nichols, P.D. y Williams, N. (2009). Consequences of test score use as validity evidence: roles and responsibilities. *Educational Measurement: Issues and Practice*, (Spring, pp. 3-9).
- Rodriguez, M.C. (2002). Choosing an item format. En G. Tindal y T. M. Haladyna (Eds.), *Large-scale assessment programs for all students*. (pp. 213-231). Mahwah: Erlbaum.
- Scriven, M. (1967). *The methodology of evaluation* (Vol. 1). Washington: AERA.
- Supreme Education Council. (2007). Qatar Comprehensive Educational Assessment (QCEA). *QCEA 2007 Results*. Disponible en: http://www.english.education.gov.qa/section/sec/evaluation_institute/sao/_qcea

LA EXPERIENCIA CHILENA DE INSTALACIÓN DE ESTÁNDARES DE APRENDIZAJE Y LA NUEVA INSTITUCIONALIDAD PARA LA CALIDAD DE LA EDUCACIÓN

Pedro Montt

Revista Iberoamericana de Evaluación Educativa 2009 - Volumen 2, Número 1

<http://www.rinace.net/riee/numeros/vol2-num1/art2.pdf>

Fecha de recepción: 30 de marzo de 2009

Fecha de comunicación de dictamen: 25 de abril de 2009

Fecha de aceptación: 25 de abril de 2009

Chile está al inicio de una nueva etapa en sus políticas educacionales. El país por mucho tiempo puso su empeño en mejorar la calidad de la educación, dotando de recursos materiales y mejorando las capacidades docentes, pero manteniendo invariantes los componentes organizacionales y de financiamiento introducidos en los 80. Si bien ello ha permitido asegurar gobernabilidad y la obtención de muchos logros, hoy ello ya se hace insostenible si se quiere avanzar en la búsqueda de una calidad para todos. Así, al aprobar recientemente una nueva ley marco, la Ley General de Educación, se ha avanzado en la construcción de nuevos consensos buscando mover la frontera respecto de las actuales condiciones institucionales en que opera el sistema escolar, abriendo una nueva etapa, la de la calidad.

La idea fuerza de las reformas de los '80 era que la subvención, la entrada de proveedores privados y la libre entrada y salida de colegios, serían los principales mecanismos que asegurarían el funcionamiento eficiente del sistema. Los gobiernos de la Concertación, mantuvieron la estructura y organización del sistema educativo y optaron por incrementar y mejorar los insumos educativos. Los sucesivos gobiernos actuaron sobre múltiples factores para mejorar el sistema educacional y como consecuencia del aumento de recursos y de las políticas seguidas a partir de los 90, ha habido una importante expansión de las oportunidades educativas de nuestra niñez y juventud.

Chile es otro en educación, pero los mismos avances logrados han hecho más visibles las inequidades existentes y han abierto nuevos desafíos. Los resultados de aprendizaje de los alumnos interpelan, hay insatisfacción y convencimiento de que se debe lograr mucho más y, a la vez, disminuir las brechas de esos resultados, en los que se evidencia una gran incidencia del nivel socioeconómico de las familias de los alumnos.

Chile puede ingresar hoy en una revisión de los dispositivos de calidad del sistema, porque muchos de los restantes objetivos han sido alcanzados. Entre otros: altas tasas de escolarización, aumento del tiempo de escolaridad de sus alumnos, inclusión de niños de los sectores más desfavorecidos; generación de dispositivos compensatorios para escuelas que atienden a poblaciones escolares de mayor vulnerabilidad; establecimiento de un currículo actualizado; profesionalización de su cuerpo docente; existencia de mecanismos de evaluación de sus docentes; incorporación de innovaciones y estrechamiento de la brecha digital y un significativo esfuerzo por aumentar el presupuesto para educación. No obstante ello, es necesario reconocer que nos encontramos en un punto de inflexión, ante el cual se deben realizar correcciones importantes si queremos satisfacer la demanda por calidad de los jóvenes, de sus familias, de Chile en general.

No podemos dejar de reconocer también, que los paros, las protestas, en parte son expresión de un nivel de frustración –de dimensión societal- respecto de los niveles de desigualdad que existen en cuanto a resultados educativos y que el sistema escolar no logra aún superar. Este es el desafío. No solo lograr una educación para todos, sino de una educación de calidad para todos.

Una tendencia internacional para avanzar en el logro de mejores resultados de aprendizajes, a la cual se ha sumado Chile, para una mejor y mayor calidad en su educación, es la instalación de estándares, entendiendo que un estándar –como lo define el PREAL (1999)- en términos generales, *“es tanto una meta (lo que debiera hacerse) como una medida de progresión hacia esa meta (cuán bien fue hecho)”*.

El desarrollo de estándares, es sostiene Margaret Forster (2007)¹, "el primer paso para impulsar el cambio. El currículum se modifica para explicitar de mejor manera los resultados de aprendizaje que se espera de todos los alumnos (estándares de 'contenido') y se desarrollan estándares de 'desempeño' para aclarar el desempeño que se espera de los alumnos en diferentes etapas de su escolaridad. En muchos países [como sería el caso de Chile], se emplean pruebas alineadas a estándares explícitos para obtener evidencia del desempeño de los alumnos y para que sirvan como base para los sistemas de rendición de cuentas (*accountability*)" (p. 2)

Sabiendo que la calidad varía con el tiempo, al responder a los cambios de la sociedad, resulta necesario acordar los diversos ámbitos o áreas en los que se debe poner especial atención para asegurar la calidad; como también, establecer los elementos que permitan discriminar entre quiénes están más cerca o más lejos de lo que se ha establecido como óptimo. Un valor instrumental importante de los estándares, como sostiene Diane Ravitch (1996), es que "ofrece una perspectiva de evaluación realista; si no hubiera modo de saber si alguien está en realidad cumpliendo con el estándar, no tendría ningún valor o sentido. Por lo tanto, cada estándar real está sujeto a observación, evaluación y medición"

Ciertamente, el establecimiento e implementación de estándares no es una tarea fácil, incluso en los casos en cuales existe una fuerte demanda y apoyo en favor de su determinación. Sin embargo, la ausencia de estándares compartidos y aceptados impide basar en ellos las pruebas de rendimiento y hace difícil interpretar los resultados de los sistemas de evaluación, como asimismo medir avances en el logro de los objetivos acordados. Sin estándares resulta más difícil y complejo responsabilizar a los centros de formación por el aprendizaje de los estudiantes y por tanto garantizar que la educación cumpla con la pertinencia y calidad necesaria.

En educación, el término "estándar" significa distintas cosas para diferentes personas. Algunas veces, por falta de precisión, se cree tener estándares cuando lo que en realidad se tiene son estímulos o confusas afirmaciones sobre aspiraciones prácticamente inmensurables. En otras ocasiones, se usan términos como "estándares", "resultados" y "metas" alternativamente, sin definir ningún significado en particular.

Siguiendo a la autora ya citada, Diane Ravitch (1996), el término estándar tiene tres usos comunes, cada uno con un propósito y significado distinto. Estos son:

- **Estándares de contenido (o estándares curriculares).** Estos estándares describen lo que los profesores debieran enseñar y lo que se espera que los estudiantes aprendan. Ellos proporcionan descripciones claras y específicas de los conocimientos que deben aprender los estudiantes. Todos los involucrados en el proceso debieran tener fácil acceso a un programa con estándares de contenido, de tal modo que las expectativas sean bien comprendidas. Un estándar de contenido debiera, además, ser medible para que los estudiantes puedan demostrar su dominio de destrezas o conocimientos.
- **Estándares de desempeño escolar.** Los estándares de desempeño definen grados de dominio o niveles de logro. Ellos responden a la pregunta "¿Cuán bueno es lo suficientemente bueno?". Los estándares de desempeño describen qué clase de desempeño representa un logro inadecuado, aceptable o sobresaliente. Los estándares de desempeño bien diseñados indican

¹ La Dra. Margaret Forster fue la responsable de la asesoría que el Ministerio de Educación de Chile solicitó al Australian Council for Educational Research (ACER) para desarrollar estándares de desempeño.

tanto la naturaleza de las evidencias (tales como un ensayo, una prueba matemática, un experimento científico, un proyecto, un examen, o una combinación de éstos) requeridas para demostrar que los estudiantes han dominado el material estipulado por los estándares de contenido, como la calidad del desempeño del estudiante (es decir, una especie de sistema de calificaciones).

- **Los Estándares de oportunidad para aprender, o transferencia escolar** definen la disponibilidad de programas, el personal, y otros recursos que la institución proporciona para que los estudiantes puedan ser capaces de satisfacer estándares de contenido y de desempeño desafiantes.

Estos tres tipos de estándares están interrelacionados. No tiene sentido contar con estándares de contenido sin estándares de desempeño escolar. Los estándares de contenido definen qué debe ser enseñado y aprendido; los estándares de desempeño describen cuán bien ha sido aprendido. Sin estándares de contenido y de desempeño, no hay forma de determinar objetivamente si el despliegue de recursos ha sido efectivo.

A esta tipología, es necesario agregar una cuarta distinción: el **estándar de desempeño docente**, que posibilita el logro de los estándares de contenido. Describe lo que los docentes deben saber y saber hacer; es decir lo que deben aprender y cómo deben desempeñarse. Esta cuarta categoría permite dar cuenta y comprender lo que en Chile se conoce como Estándares para el desempeño docente –Marco de la Buena Enseñanza– que operan para la evaluación de los docentes en servicio en los establecimientos educacionales públicos. En un futuro cercano este marco se complementará con la existencia de estándares para la formación inicial docente.

1. CONTEXTO Y ORIGEN DE LOS ESTÁNDARES EN CHILE

En Chile la iniciativa de desarrollar estándares se inscribió en el contexto de una preocupación general por incidir mayormente en la calidad y equidad educacional.

Haciendo un recorrido histórico en estas materias, Jacqueline Gysling y Lorena Meckes (2007)², establecen los siguientes hitos. El año 2000 se publicaron los resultados del SIMCE³ de 4° básico aplicado el año 1999, y ese mismo año los resultados de octavo básico TIMSS. Ambos resultados generaron un gran impacto en la opinión pública y en el Ministerio de Educación. En el caso del SIMCE, porque por primera vez se describió niveles de desempeño y se determinó el alto porcentaje de niños y niñas que no alcanzaba en 4° básico aprendizajes que se esperaban para 2° básico según el currículum. Por otra parte, TIMSS mostró la distancia entre los aprendizajes logrados por los estudiantes chilenos, incluidos los de la

² En los tres primeros apartados: a) Contexto y Origen de los estándares en Chile, b) Características de los estándares en Chile y c) Nudos centrales durante el proceso de instalación de estándares; se utiliza como base, con la autorización de las autoras, en forma bastante textual, el trabajo de Gysling, Jacqueline y Meckes, Lorena (2007) " *Mapas de Progreso y Niveles de Logro SIMCE. Desarrollo de Estándares de Contenido y Desempeño en Chile: 2002 a 2007* ". Documento de trabajo interno de la Unidad de Currículum y Evaluación del MINEDUC de Chile.

³ El SIMCE es el sistema nacional de medición de resultados de aprendizaje del Ministerio de Educación de Chile. Su propósito es contribuir al mejoramiento de la calidad y equidad de la educación, informando sobre el desempeño de los alumnos y alumnas en diferentes áreas del currículum nacional, y relacionándolos con el contexto escolar y social en el que ellos aprenden. Las pruebas SIMCE evalúan el logro de los Objetivos Fundamentales y Contenidos Mínimos Obligatorios del Marco Curricular en diferentes subsectores de aprendizaje, a través de una medición que se aplica a nivel nacional, una vez al año, a los estudiantes que cursan un determinado nivel educacional. A partir del año 2006, se evalúa todos los años a 4° Básico y se alternan 8° Básico y 2° Medio. Para mayor información ver www.simce.cl

elite, y los alcanzados por los países desarrollados. La prensa abordó esta información con una ola de críticas a la reforma desarrollada en los 90, y se consolidó la necesidad de revisar la política educativa implementada por los Gobiernos de la Concertación.

Así, se reorientó la política con el propósito de tener mayor incidencia en los resultados de aprendizaje: bajo el lema “llevar la reforma al aula” se formularon orientaciones curriculares más detalladas de primero a cuarto básico y se inició una campaña de Lectura Escritura y Matemática (LEM). Como parte de esta campaña, se elaboró un documento de orientación para las escuelas, en que se describía, ejemplificaba e ilustraba los “desempeños a lograr” en 2º nivel de transición, 2º y 4º básico para Lectura, Escritura y Matemática (MINEDUC 2003a) siguiendo una estructura muy similar a la que tendrían posteriormente los estándares. Además, se intervino con asesoría externa (provista por ONGs y Universidades) a las escuelas con peores resultados de la Región Metropolitana. Los distintos programas de mejoramiento ministeriales se orientaron a implementar el currículum. Paralelamente, el año 2002 la Unidad de Currículum y Evaluación del Ministerio de Educación inició la elaboración de estándares que se consideraban un elemento básico del nuevo énfasis en los resultados de aprendizaje.

Adicionalmente, el Ministerio solicitó a la OECD hacer una evaluación del conjunto de la reforma educacional de los noventa y convocó a una Comisión de expertos y actores educativos para que revisara el SIMCE e hiciera recomendaciones para su mejoramiento. Ambas comisiones operaron el año 2003, y entre sus recomendaciones reforzaron la necesidad de formular estándares para otorgar significado a la medición nacional y señalar con mayor claridad la expectativa de resultados en ella (Mineduc, 2003b).

La Comisión de la OECD planteó esta necesidad en el contexto de una crítica mayor acerca del ordenamiento institucional del sistema educativo chileno en un doble sentido. Por una parte una excesiva confianza en mecanismos de mercado e incentivos como forma de regulación, y por la otra una descentralización incompleta especialmente respecto de la educación municipalizada (OCDE, 2004a). Estos problemas estructurales se retomaron con fuerza y gran impacto social con el movimiento estudiantil del año 2006, producto de este movimiento, se convoca a un Consejo Asesor Presidencial integrado por representantes estudiantiles, del Colegio de Profesores, de los padres de familia, de los distintos actores del sistema educacional y diversos expertos en materia educacional, incluyendo a varias de las personas que encabezaron el proceso de reforma educacional de los noventa. Este Consejo generó un conjunto de recomendaciones de diversa índole, entre las cuales se incluyó estándares de aprendizaje y estándares de calidad de los centros educativos, junto con la generación de una entidad estatal que fiscalice el cumplimiento de ellos.

En este marco y frente a estas demandas, el Gobierno respondió, en primer lugar, enviando al Congreso una reforma a la Constitución estableciendo el derecho a una educación de calidad, lo que consagra la obligación del Estado de velar por ello. Además, se ha aprobado una Ley General de Educación y se encuentra avanzado el trámite de un proyecto que crea una Superintendencia y una Agencia de la Calidad de la Educación. En esta ley ya aprobada y el proyecto de Aseguramiento de la Calidad (Superintendencia y Agencia) tanto en la postura del gobierno como de la oposición política, los estándares de aprendizaje juegan un rol articulador de la organización del sistema.

2. CARACTERÍSTICAS DE LOS ESTÁNDARES EN CHILE

En Chile, los estándares de aprendizaje que ha elaborado el Ministerio de Educación distinguen entre *estándares de contenido* y *estándares de desempeño*. Los primeros descritos como Mapas de Progreso del Aprendizaje en los ejes centrales de cada área curricular, y los segundos -denominados Niveles de Logro- como descripciones de niveles crecientes de desempeño en determinados momentos de la vida escolar de los estudiantes, específicamente para las áreas y los grados que son monitoreados por la medición nacional SIMCE.

Tanto los niveles de los mapas, como los Niveles de Logro del SIMCE describen cualitativamente el desempeño mostrado por los estudiantes. En el caso de los Niveles de Logro o estándares de desempeño, se operacionalizan además cuantitativamente en un puntaje de corte en la medición nacional, que corresponde al puntaje mínimo que debe obtener un estudiante en la prueba, para que su desempeño sea clasificado en el nivel o categoría de rendimiento respectiva.

Las principales características y opciones de enfoque que adoptó el desarrollo de estándares en Chile son:

1. Describir el aprendizaje como progresión
2. Basar la elaboración de los estándares en evidencia empírica
3. Articular bajo un mismo marco de referencia la medición nacional y el monitoreo del aprendizaje que se realiza a nivel local y
4. Otorgar preponderancia a la comunicación y uso en la elaboración de estándares.

2.1. Describir el aprendizaje como progresión

Tanto los Niveles de Logro del SIMCE como los Mapas de Progreso del aprendizaje se basan en el supuesto que el aprendizaje es un continuo que se enriquece a lo largo de la trayectoria escolar. Desde esta perspectiva el aprendizaje no es una sumatoria de conocimientos que se van adquiriendo en forma discreta o aislada, como capas de información, sino que el aprendizaje, desde el punto de vista del sujeto, es un proceso de desarrollo de capacidades, entendimientos, habilidades de realización que se van profundizando y ampliando desde niveles más simples a niveles más complejos. Esta noción del aprendizaje está estrechamente vinculada a la noción de competencia, entendida como un sistema complejo de acción que integra conocimientos, habilidades, actitudes, valores que se ponen en juego en contextos reales de ejecución.

Los Mapas de Progreso no son simples listas de resultados específicos como era el caso en los esquemas curriculares basados en objetivos que se usaron a partir de fines de los años sesenta. Por el contrario, los niveles de los Mapas de Progreso son descripciones amplias del aprendizaje que permiten a los docentes realizar juicios usando diversas evidencias, sobre el nivel en que se encuentran sus alumnos.

Como sostiene Margaret Forster (2007) "es fácil suponer que los docentes comprenden lo que significa que un alumno progrese en un área de aprendizaje. Sin embargo, la realidad es que a muchos docentes se les dificulta proporcionarles a los alumnos experiencias de aprendizaje apropiadas y retroalimentación que les ayude a avanzar en su aprendizaje" (p. 2). De aquí, sostiene Forster, la importancia de *mapear* "el crecimiento explícitamente con el propósito de proporcionarle un marco y una herramienta a los docentes" (p. 2). "Una vez que el docente comprende la naturaleza del crecimiento en un área de

aprendizaje y ha establecido el nivel actual de logros del alumno –afirma la misma autora- podrá tomar decisiones, basadas en evidencias, acerca de cómo facilitar de mejor manera un mayor aprendizaje” (p. 2)

En el caso de los Niveles de Logro, estas descripciones son más específicas y precisas ya que se espera diferenciar niveles crecientes de dominio del área evaluada dentro de un mismo grado (por ejemplo diferenciar niveles de logro en 4° grado).

Tanto en los Mapas de Progreso o *estándares de contenido* como en los Niveles de Logro o *estándares de desempeño*, cada nivel es inclusivo del anterior, y representa un grado mayor de desarrollo o complejidad de las tareas que el estudiante demuestra que es capaz de realizar.

Es importante señalar que los Mapas de Progreso y los Niveles de Logro parten del supuesto que en un mismo curso (grado, aula) se encontrarán alumnos en distintos niveles del mapa. Los profesores son concientes de esta diversidad, pero suelen ordenarla definiendo un grupo de buenos alumnos, otro de alumnos regulares y otro de malos alumnos en base a sus propios criterios o parámetros. Lo que se pretende con este instrumento es que los profesores describan la diversidad a partir del mapa, para entender que sus alumnos se encuentran en un determinado nivel de aprendizaje, que el docente debe conocer para conducir a su alumno a su nivel de desarrollo próximo. Describir el aprendizaje en progresión, es muy relevante desde el punto de vista del mejoramiento, ya que lo que se busca no solo es describir si un alumno ha alcanzado o no determinado nivel, sino poder describir dónde está en relación al estándar y cuán lejos está de alcanzarlo. Este es el mismo concepto que está tras los Niveles de Logro que describe el SIMCE.

Siguiendo estas consideraciones los mapas combinan una expectativa del aprendizaje ideal para determinado grado escolar, con la posibilidad de describir para ese grado la diversidad de niveles de logro que puede existir en ese grupo de niños y niñas. Por esto, los niveles de los mapas tienen una asociación a determinados grados escolares. Esta asociación es un mensaje curricular importante, ya que define el logro esperado en un determinado nivel, además de ser un mecanismo de establecer coherencia interna a la construcción, tanto en sentido vertical como horizontal entre mapas. Por ejemplo, todos los niveles 1 de los mapas están asociados al segundo básico (o sea, 2do grado de la escolaridad), es decir describen el aprendizaje que es típico en cada eje a los 7 años de edad. En el caso chileno la asociación a grados se ordena de dos en dos, es decir cada nivel del mapa de progreso representa la expectativa de aprendizaje de dos años de trabajo escolar; el nivel 7, el más alto, representa un nivel de logro que supera la expectativa del último año escolar en educación media (4to medio).

2.2. Basar la elaboración de los estándares en evidencia empírica

En Chile, tanto la elaboración de los Mapas de Progreso como de los Niveles de Logro recurrió al análisis de evidencia sobre resultados en mediciones nacionales e internacionales. La elaboración implicó el constante diálogo entre el juicio experto sobre qué resulta logable en determinadas etapas del trayecto escolar, sobre qué aprendizajes resultan más o menos complejos desde el punto de vista de la disciplina, con la evidencia acumulada y disponible a partir de pruebas internacionales y de la administración de pruebas SIMCE desde 1999.

En específico, se revisaron tanto las descripciones de logro reportadas por las pruebas internacionales para distintos tramos de rendimiento como preguntas específicas y sus respectivos datos sobre el logro alcanzado para cada una de ellas por parte de estudiantes chilenos cuando este dato estaba disponible

(SIMCE y pruebas internacionales en que Chile ha participado), como por estudiantes de los distintos países si se trataba de mediciones nacionales o estatales.

Esta revisión de evidencia permitió tanto para Niveles de Logro como para Mapas de Progreso:

- examinar qué se valoraba (incluía) en las distintas mediciones nacionales e internacionales y sus respectivos estándares o descripciones cualitativas del logro alcanzado para contrastar los primeros borradores de los Mapas de Progreso y las dimensiones que se considerarían en los Niveles de Logro y tipos de pregunta que sería interesante incorporar en las pruebas SIMCE,
- ajustar el nivel de exigencia tanto de los estándares de contenido como de desempeño considerando tanto las descripciones generales como los indicadores,
- alinear el nivel de exigencia entre las distintas áreas (por ejemplo entre Matemática y Lectura), y
- afinar lo que se ha llamado “identidad” del nivel u homogeneidad de su dificultad procurando que en un mismo nivel no se incluyeran logros que de acuerdo con lo que muestran las pruebas internacionales “pertenecen” a niveles de dificultad diferentes.

2.3. Articular, bajo un mismo marco de referencia, la medición nacional y el monitoreo del aprendizaje que se realiza a nivel local

En Chile, como describe Margaret Forster (2007), “la intención fue ubicar los estándares de desempeño dentro del mismo marco conceptual que los estándares de contenido. Así, estos podrían constituirse en marcadores [*markers*] a lo largo del continuo que describe el mapa de progreso. La necesidad de contar con un marco conceptual único no es sólo una cuestión teórica. Si a los docentes se les proporcionan dos marcos teóricos con los cuales lidiar, lo más probable es que se concentren en aquel que perciben como más importante –en desmedro del otro” (p.4).

La expectativa de desarrollar estándares de desempeño y de contenido alineados entre sí y bajo un mismo enfoque, busco además, ofrecer al sistema un marco común para la medición externa y para el monitoreo de los logros de aprendizaje a nivel local (municipios, regiones y las mismas escuelas). De este modo, promueve y favorece la articulación de la información recabada a nivel local con la información entregada por la medición nacional, potenciando el uso de ambas en el diseño de planes de mejoramiento.

Este enfoque combina mecanismos de control externo (foco en áreas claves del currículum, cuenta pública y responsabilización) con mecanismos de mejoramiento (desarrollo de capacidades, gestión local, currículum integral) bajo el supuesto que la sola existencia de controles externos sin complementar con el apoyo adecuado no produce efectos sostenidos en la calidad del sistema educacional (OCDE 2005 y 2004b).

En este campo Chile ha seguido la experiencia internacional, que establece que en un sistema educacional bien integrado, **estándares y evaluación van de la mano**. Los estándares sirven como un recurso indicativo importante para los estudiantes, profesores y universidades en cuanto institución, transmitiendo a cada uno lo que se espera de ellos; mientras que las evaluaciones proporcionan información acerca de cuán bien se han cumplido las expectativas.

Según Linn y Herman (1997), la medición basada en estándares tiene las siguientes características:

- **Está ligada estrechamente al currículum:** La medición debe ser congruente con el currículum. Esta relación de congruencia debe evidenciarse en el alineamiento entre los elementos del

proceso educativo: a saber: los estándares y el currículum, la formación y capacitación docente, los materiales de enseñanza, los libros de texto y la medición del logro.

- **Compara el rendimiento de cada estudiante con las metas de contenido y ejecución previamente establecidas.** La medición basada en estándares es fundamentalmente criterial. Lo que se espera es obtener información sobre el logro específico de conocimientos, habilidades o actitudes por parte de un alumno, y no sobre el rendimiento de un alumno con respecto a su grupo, que es propio de una evaluación normativa.
- **Incorpora nuevas formas de medición del logro, conocidas como medición del desempeño y medición auténtica.** Se promueve una auténtica evaluación de desempeño que fomente un rol activo del alumno, desarrollando tareas complejas en contextos cotidianos o escenarios reales. La medición de desempeño basada en evaluación auténtica establece criterios previos mucho más pertinentes, aceptados y significativos, lo cual favorece la transferencia de la responsabilidad al estudiante, el desarrollo de habilidades de autorregulación y procesos de metacognición. Se emplean, para estos efectos, medios y estrategias que involucran al estudiante en forma protagónica, que ocupan tiempos sustantivos y que despliegan el uso de estrategias de aprendizaje de mayor elaboración.

Dentro de un sistema integrado en pos de la calidad, **los estándares están también relacionados con los programas de mejoramiento.** Evaluación, determinación de estándares y programas de mejoramiento, son etapas de un sólo proceso que permite avanzar a la calidad. Los estándares indican a los establecimientos lo que tienen que hacer para su éxito; las evaluaciones indican si están progresando en el sentido presupuestado; los programas de mejoramiento son las acciones que se emprenden para transitar desde lo que se tiene como realidad hasta lo que se espera como ideal.

No basta, por lo tanto, disponer de estándares que permitan establecer qué es lo adecuado, ni tampoco basta poseer los mecanismos de evaluación que posibiliten determinar la etapa en que se ubica una persona o institución, en relación a dicho estándar. El ciclo se cierra con la posibilidad de generar un programa de mejoramiento, que permita el paso desde lo que se tiene (lo que arroja la evaluación) a lo que se quiere (lo que se considera como producto de calidad).

El sistema ideal, como afirma Forster (2007), consiste en "componentes interconectados que se apoyan mutuamente: un mapa de progreso que constituye la columna vertebral del sistema; una gama de métodos de evaluación que permite la obtención de evidencias de los logros de los alumnos; oportunidades para actividades de desarrollo profesional con el propósito de apoyar a los docentes en el uso de estos métodos; herramientas para ayudar a los profesores a evaluar el trabajo de los alumnos; procedimientos para recopilar información confiable del sistema en su totalidad; y procesos para informar los niveles de desempeño, a nivel de escuela y de sistema, y para monitorear niveles de logros en el tiempo" (p. 6).

2.4. Otorgar preponderancia a la comunicación y uso en la elaboración de estándares

La importancia de la comunicación y el uso de los estándares se hizo evidente tanto en el énfasis puesto en la claridad y síntesis de su formulación, como en las sucesivas validaciones que se realizaron para asegurar esta claridad y facilitar su uso respecto del lenguaje utilizado, así como de los formatos gráficos en que se iba a presentar la información del SIMCE.

Para ayudar a los docentes a observar y verificar el aprendizaje de cada nivel, tanto los mapas de progreso como los estándares de desempeño de SIMCE se **ilustran** con ejemplos de desempeños (indicadores). Cada nivel de los Mapas de Progreso se ilustran además con ejemplos de trabajos realizados por los estudiantes y los estándares de desempeño se ilustran con preguntas de la prueba que los alumnos de cada nivel son capaces de responder.

Tanto *indicadores* como *ilustraciones comentadas* han tenido el propósito de clarificar la exigencia o nivel de desempeño que se describe, procurando mostrar los aprendizajes característicos o distintivos de este.

Los ejemplos de trabajo de los alumnos contestan la pregunta: ¿Cuál es la apariencia del aprendizaje en este nivel?

Dado que los ejemplos de trabajo de los alumnos se recolectan con una tarea determinada que ellos ejecutan, éstos también ilustran el tipo de actividades a realizar para observar el aprendizaje. Adicionalmente, y dependiendo del tipo de comentarios que se le hagan a los trabajos, en estos ejemplos se pueden dar orientaciones sobre cómo apoyar el progreso de los alumnos.

Los estándares de desempeño o niveles de logro presentan una misma estructura: descripción general o enunciado, indicadores e ilustraciones de preguntas comentadas. En el caso de los **Niveles de Logro**, se incorporó además un comentario de contraste entre el tipo de tareas característicos o expresivos de cada nivel, para apoyar la comprensión sobre “aquello que progresa” entre un nivel y el otro.

A partir del año 2006, el SIMCE informa qué porcentaje de alumnos alcanza distintos niveles de logro para cada escuela⁴ y para diferentes agregaciones territoriales, socioeconómicas, administrativas, etc. Esto permite monitorear las estrategias implementadas desde distintos niveles de administración según el nivel de aprendizajes que los alumnos están logrando. A partir también de esa fecha, las escuelas conocen cuántos de sus alumnos de 4° básico han alcanzado cada nivel de logro y cómo se distribuyen los alumnos al interior de cada nivel.

La instalación de ambos tipos de estándares ha seguido el calendario que se muestra a continuación:

TABLA 1.

Estándares	2006	2007	2008	2009	2010	2011	2012
Mapas de Progreso		5*		8**	6		
Niveles de Logro 4° grado	Lenguaje Matemática	Ciencias Naturales	Ciencias Sociales				
Niveles de Logro 8° grado				Lenguaje Matemática		Ciencias Naturales y Ciencias Sociales	
Niveles de Logro 10° grado							Lenguaje Matemática

*Uno por sector curricular ajustado: Lenguaje; Matemática; Ciencias Naturales; Ciencias Sociales e Inglés.

**Se agregan mapas a cada sector, ya que éstos constituyen ejes. Por ejemplo Lenguaje posee tres mapas: Lectura, Escritura y Comprensión Lectora.

FUENTE: Unidad de Curriculum y Evaluación, Mineduc, Chile.

⁴ Hasta el 2006 el SIMCE había informado públicamente un puntaje promedio, haciendo invisible que al interior de cada escuela hay una gran dispersión, y que en un mismo establecimiento hay alumnos que alcanzan muy altos puntajes y otros logran resultados extremadamente bajos. Si bien desde 2005 se había entregado a cada establecimiento la dispersión de sus puntajes, esta variación interna no era objeto de publicación ni discusión a nivel nacional.

3. NUDOS CENTRALES DURANTE EL PROCESO DE INSTALACIÓN DE ESTÁNDARES DE APRENDIZAJE

La tarea de instalar estándares no es exenta de grandes desafíos y tensiones que implican tomar decisiones en uno u otro sentido. Al respecto, es relevante destacar los siguientes nudos críticos.

3.1. El difícil complemento y balance entre presión y apoyo

La tensión entre mecanismos de presión o de apoyo se ha manifestado en diversas formas a lo largo del proceso de elaboración de estándares de aprendizaje y la respuesta a la pregunta sobre cuál será el que predomine está sujeta al predominio que alcance una u otra visión. Esta búsqueda de equilibrio entre ambos tipos de uso de los estándares se ha traducido en la práctica en decisiones de comunicación y de articulación de diversas estrategias. Por ejemplo, en la perspectiva de control, una línea de implementación es la restricción de la autonomía de las escuelas en el uso de recursos adicionales de acuerdo a resultados en el SIMCE, o la evaluación de desempeño de los directores municipales de acuerdo al cumplimiento de metas de mejoramiento asociado a resultados en pruebas nacionales o regionales. Por otro lado, el uso como apoyo a la gestión y a la pedagogía, se materializa a través de la promoción de instancias de reflexión entre docentes y administradores sobre el aprendizaje de los alumnos y los modos de observarlo en el aula y en la escuela, usando para ello tanto los mapas de progreso como los resultados del SIMCE con niveles de logro.

Esta tensión entre presión y apoyo y el esfuerzo puesto en prevenir resistencias ha permeado múltiples decisiones del desarrollo de los estándares, desde el mismo modelo que se ha adoptado para su elaboración, como decisiones respecto a la comunicación. Por ejemplo, sustituir sistemáticamente los términos “estándares de contenido” y “estándares de desempeño” por Mapas de Progreso y Niveles de Logro respectivamente, ha obedecido a un esfuerzo deliberado por reducir posibles connotaciones de control o uniformidad, que típicamente evoca el término “estándar”.

Otra manifestación de esta tensión se presentó ante la decisión de continuar reportando públicamente (en el periódico y en la web) sólo los puntajes promedio por escuela y no la distribución de alumnos según estándares de desempeño para mitigar el impacto que pudiera tener ser expuestos y comparados públicamente frente a expectativas de logro⁵. También esta tensión se ha manifestado en las ambivalencias para asociar abiertamente los Mapas de Progreso a la *evaluación del aprendizaje* e incluso al SIMCE, y la preferencia por presentarlos como un apoyo para orientar la pedagogía. En esta misma línea el énfasis de la comunicación pública se ha puesto en que estas herramientas permiten evidenciar la diversidad de aprendizajes y la relevancia de esto para conducir mejores prácticas docentes por sobre su potencial de uso para control en torno a una expectativa de logro. Al mismo tiempo, se ha promovido los usos de estándares en las leyes que definen la nueva institucionalidad del sistema, como parámetros para ejercer un control central sobre la calidad.

3.2. La preocupación por la coherencia sistémica: relación entre currículo, prueba, estándares de desempeño y de contenido

Los estándares de contenido y de desempeño se han elaborado articuladamente, tanto en términos institucionales como conceptuales, y se ha previsto asegurar la alineación con el currículum vigente, a la

⁵ Si bien el SIMCE publica los puntajes promedio por escuela desde 1995, se decidió entregar sólo a cada comunidad educativa (sus profesores y padres) la distribución de los alumnos según los estándares de desempeño sin hacerlos públicos en la web o el diario, al menos por la primera vez.

vez que alcanzar una mayor especificación de este. En términos institucionales, los profesionales a cargo de la elaboración, pertenecen a la misma organización (Unidad de Currículo y Evaluación) y corresponden a las unidades de currículum y SIMCE. Estos equipos se hicieron cargo respectivamente de los mapas de progreso y de los estándares de desempeño y su medición. Los estándares de contenido se han elaborado como un complemento al marco curricular y a los programas de estudio para reforzar la evaluación del aprendizaje que promueve el currículum.

No obstante las precauciones adoptadas, la alineación entre la prueba, niveles de logro, currículum y mapas de progreso ha sido un desafío permanente, procurando que los mensajes curriculares que reciba el sistema escolar sean coherentes entre sí para evitar la confusión. En este campo la tensión es que los estándares de contenido se mantengan en el enfoque curricular vigente, pero muchas veces en su elaboración se hace evidente la necesidad de ajuste de la secuencia curricular. A su vez el SIMCE, que por disposición legal debe evaluar el currículum vigente, no incorpora a la medición –con igual rapidez- las innovaciones al currículum que los mapas ya anticipan.

Las ventajas de la articulación institucional existente a la fecha, han sido el uso fluido de evidencia proveniente de las pruebas para reformular permanentemente el currículum y elaborar los estándares, tanto de contenido como de desempeño y la posibilidad de anticiparse a los ajustes curriculares e incorporar las novedades en las pruebas experimentales de SIMCE dos años antes, para así evitar desfases entre la medición y el currículum vigente, entre otros. Todo lo anterior, considerando la relevancia de la coherencia entre mensajes que se entrega al sistema sobre expectativas de aprendizaje.

3.3. ¿Cuán basados en evidencia?

Entre las características de los estándares se incluyó su base en evidencia empírica, tanto para decidir si determinados descriptores o indicadores correspondían a un determinado nivel como para evaluar cuán “homogénea” era la dificultad de cada nivel.

En el caso de los estándares de desempeño se buscó ser muy rigurosos, incluyendo en las pruebas piloto llamadas “de evidencia” preguntas para indagar su nivel de complejidad para los niños chilenos de 4º básico, antes de incluirlas entre los indicadores o en la descripción. Esto implicaba formular hipótesis sobre la progresión del aprendizaje, es decir hipótesis sobre en qué consiste avanzar en un mismo grado o curso, por ejemplo, en la capacidad para reflexionar u opinar sobre el texto leído. En diferentes oportunidades las hipótesis no se confirmaban, sin embargo, en ocasiones era imprescindible incluir la dimensión que se estaba explorando en la descripción o en los indicadores aun cuando se contara con muy pocos datos para fundamentar el indicador o descripción. Así, en estos casos se complementaba la descripción “a posteriori” de aquello que efectivamente mostraban los estudiantes en las pruebas con una visión “a priori” o teórica sobre aquello que era relevante incluir en la descripción. En este sentido, una pregunta permanente en la elaboración de estándares de desempeño es ¿cuánta evidencia (número de preguntas de la prueba que se sitúan en “la zona” que se está describiendo) es necesaria para poder fundamentar una descripción?

En los casos en que no se puede tomar como punto de partida los resultados en pruebas anteriores -al no estar presentes hasta ese momento entre los objetivos del currículum, por lo tanto nunca han sido evaluados por la prueba SIMCE y por ello no se cuenta con evidencia respecto del rendimiento de los estudiantes- se necesita tomar como punto de partida una descripción “a priori” basada en la expectativa señalada en los mapas de progreso y el nuevo currículum, para luego corroborarla con la evidencia de la

prueba experimental. En esta oportunidad se hace necesario recurrir a descripciones (benchmarks) de pruebas a nivel internacional, para verificar el progreso entre un nivel y otro de los estándares de desempeño, y la medida en que se está llegando a descripciones "homogéneas en dificultad".

En el caso de los mapas de progreso, la mayor dificultad ha sido que el respaldo en evidencia es desigual entre distintos mapas y en algunos casos claramente insuficientes. Típicamente las pruebas nacionales e internacionales evalúan determinados grados quedando algunos vacíos entre niveles de los mapas de progreso. Más evidente aun es la ausencia de datos para mapas tales como expresión oral, por ejemplo. Por otra parte, la determinación del nivel de exigencia haciéndolo equivalente al promedio de la OCDE es bastante estimativo.

3.4. La preparación de los equipos profesionales

La preparación de los equipos fue y seguirá siendo un desafío. Probablemente habría sido deseable una capacitación previa de un grupo que liderara luego la elaboración, y evitar o minimizar el aprendizaje vía ensayo y error. Al mismo tiempo, aun se verifica una carencia de un plan sistemático de transferencia interna del enfoque y de la reflexión que se ha ido desarrollando, la cual se realiza sobre la marcha y se distribuye desigualmente entre los miembros de los equipos tanto en SIMCE como en Currículum.

El discurso sobre los estándares se ha ido desarrollando junto con la comunicación pública, pasando rápidamente desde la elaboración al discurso sobre uso, y por ejemplo, del uso pedagógico al uso para gestión y políticas.

4. CAMBIOS EN LA ARQUITECTURA DEL SISTEMA EDUCACIONAL

Chile hoy trabaja en la instalación de una nueva arquitectura, en una nueva institucionalidad y un nuevo marco regulatorio de su educación. La agenda de la política educacional actual está traspasada por la tarea de "*levantar las vigas centrales*" del edificio, en tres ámbitos fundamentales: (a) Un nuevo Marco Regulatorio; (b) la generación de dispositivos que asocien responsabilidad y resultados y (c) el establecimiento de un modelo de gestión pública de la educación, que fortalezca la descentralización, pero que vele por resultados de calidad.

Un Nuevo Marco Regulatorio que asegure el derecho a una Educación de Calidad para Todos. En este marco se ubica una nueva ley General de Educación, que balancee adecuadamente el derecho a una educación de calidad con el principio de libertad de enseñanza y la creación de un sistema de aseguramiento de la calidad de la educación, en el cual jueguen su papel las agencias de la alta dirección sectorial: Ministerio, Consejo Nacional de Educación y una Superintendencia.

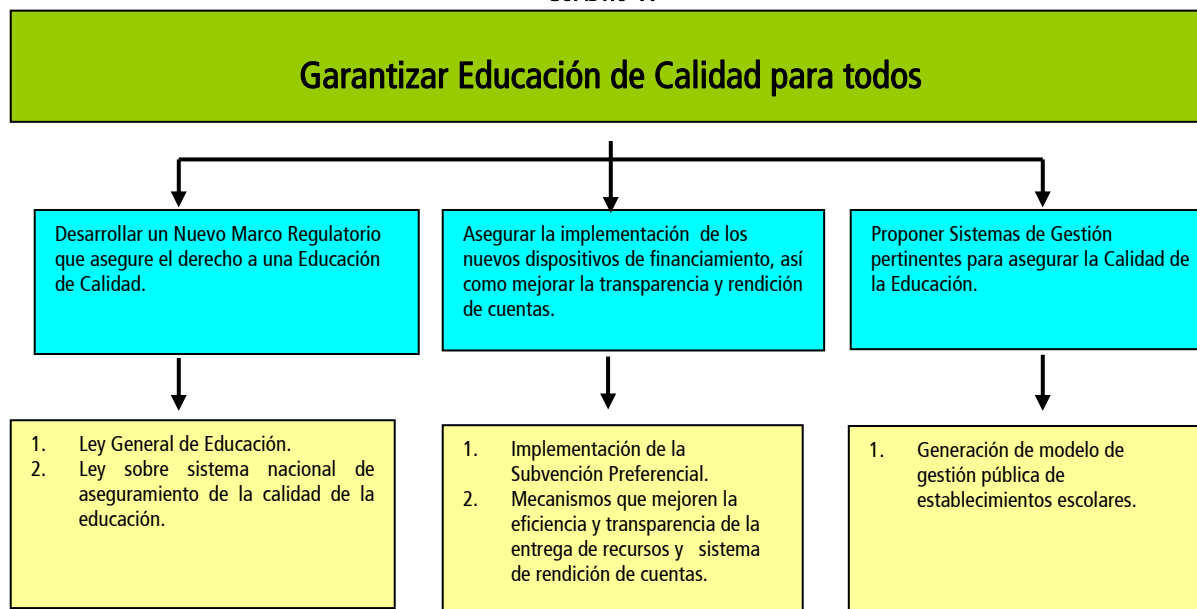
Mejorar el sistema de financiamiento de manera que permita el acceso de todos los alumnos a una educación de calidad, compense desventajas sociales y territoriales, y promueva el mejoramiento continuo del sistema educativo. En éste ámbito se ubica la implementación de la subvención escolar preferencial para la igualdad de oportunidades de niños y niñas de sectores más vulnerables, mejorar sustantivamente la transparencia y rendición de cuentas del uso de recursos públicos y, por último, asegurar que los recursos entregados por el estado estén asociados al cumplimiento de estándares mínimos de calidad del servicio educativo.

A lo antes señalados, se debe agregar el esfuerzo de fortalecer la institucionalidad y la gestión de la educación pública. En éste campo la tarea mayor se relaciona con el establecimiento de nuevos modelos

de gestión para la educación pública, que fortalezcan la descentralización, pero al mismo tiempo velen por resultados de calidad.

Ello esquemáticamente se puede representar de la siguiente forma:

CUADRO 1.



FUENTE: Elaboración propia

Esta nueva arquitectura agrega a los nudos y tensiones propias de la instalación de estándares otros nuevos desafíos, ahora en el campo mayor de la nueva institucionalidad para el sector educación.

4.1. La Ley General de Educación

Equilibrar el principio de libertad de enseñanza y derecho a una educación de calidad es uno de los objetivos de la **Ley General de Educación (LGE)**, que sustituye, en lo referido al sistema escolar, a la Ley Orgánica Constitucional de Enseñanza (LOCE). Ley, esta última, que fue aprobada en los últimos momentos del gobierno militar, sobre la cual no se realizó discusión alguna, y que es una norma que ha sido difícil de modificar por las exigencias de quórum que trae consigo.

Hoy velar por una educación de calidad, suma a las antiguas exigencias, el deber de cumplir con determinados *estándares*, que le permitan al país saber si los establecimientos educacionales caminan en el sentido correcto de la calidad para todos. De aquí que Ley General de Educación establece la existencia de un organismo, el **Consejo Nacional de Educación** (que asegura la máxima representación de la comunidad nacional), que determina los estándares de calidad que el país exigirá en educación. Esta misma exigencia de cumplir los estándares, lleva consigo, además, que junto al derecho preferente de los padres y al deber de educar a sus hijos, resulte también imperioso, consagrar el derecho que poseen las familias a ser informados sobre los resultados académicos obtenidos por sus hijos.

4.2. Proyecto de ley de aseguramiento de la calidad de la educación

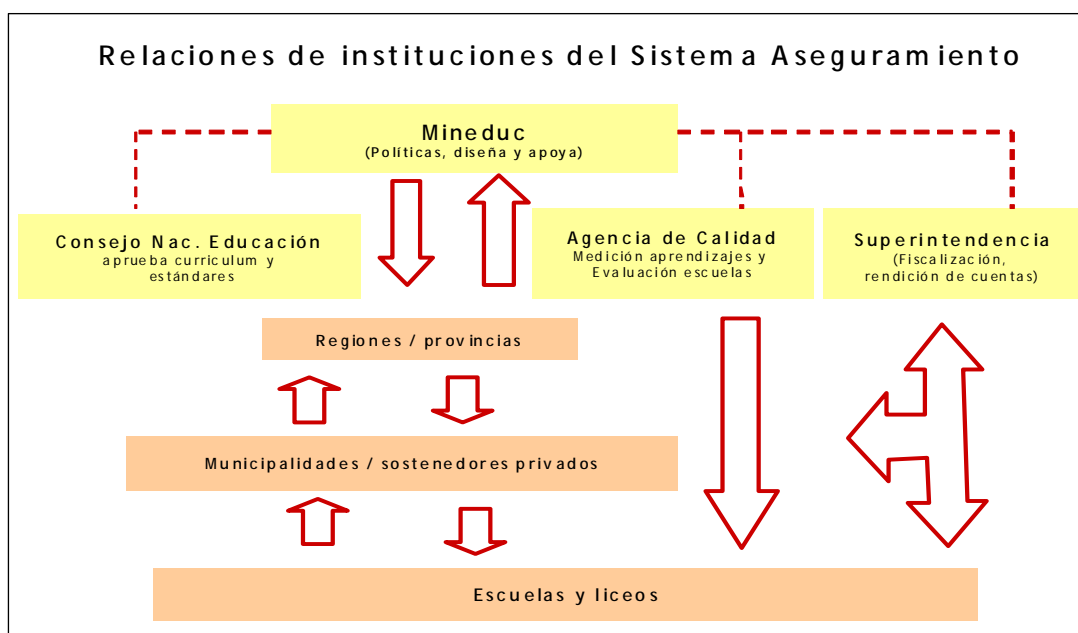
La Ley General de Educación, crea dos nuevas instituciones (que son objeto de otra ley): la **Agencia de la Calidad** y la **Superintendencia**. La creación de ambas se enmarca en el objetivo mayor de desarrollo de un Sistema de Aseguramiento de la Calidad.

La Agencia de Calidad de la Educación tendrá como objeto evaluar la calidad de la educación, para el cual cumplirá con cuatro funciones: (a) evaluar los logros de aprendizaje de los alumnos de acuerdo al grado de cumplimiento de estándares; (b) realizar evaluación de desempeño de los establecimientos educacionales y sus sostenedores, de los docentes y sus directivos en base a estándares indicativos; (c) clasificar a los establecimientos según los resultados y (d) proporcionar información en materias de su competencia a la comunidad educativa y a la ciudadanía en general.

El objeto de la Superintendencia será fiscalizar que los sostenedores y los establecimientos educacionales reconocidos por el Estado, en los niveles parvulario, básico y medio, cumplan con las leyes, reglamentos e instrucciones y con el adecuado uso de los recursos. Asimismo proporcionará información a las comunidades educativas y otros usuarios e interesados, atenderá a los reclamos de éstos y establecerá las sanciones que corresponda.

Esto último exige que el marco regulatorio compuesto por las instituciones ya identificadas, deba poseer una sinergia entre sí, que favorezca la calidad educacional, como puede apreciarse en el siguiente esquema:

CUADRO 2.



FUENTE: Elaboración propia

4.3. La Subvención Escolar Preferencial

Por décadas el financiamiento del Sistema Escolar en Chile, estableció una relación entre el Estado y las Escuelas, donde la entrega de recursos se hacía sin importar los resultados que se obtenían. Hoy se ha avanzado a una gestión del financiamiento que asegura que los recursos públicos son aplicados con efectividad al aprendizaje, condicionándolos a resultados educativos objetivos, basados en *estándares nacionales de aprendizaje*, y premiando a las escuelas que desarrollen adecuadamente los talentos de sus alumnas y alumnos. De esta forma se vincula con mayor fuerza que antes recurso con calidad.

En este marco se inscriben la Subvención Educacional Preferencial y las exigencias de transparencia y rendición de cuenta a los establecimientos educacionales.

La incorporación al sistema de Subvención Escolar Preferencial, los sostenedores de establecimientos educacionales subvencionados deben suscribir y cumplir un "Convenio de Igualdad de Oportunidades y Excelencia Educativa" con el Ministerio de Educación, en el que asumen compromisos en materia de no discriminación e igualdad de oportunidades, y se comprometen a lograr resultados educativos de calidad, sostenibles en el tiempo.

El incremento de los recursos financieros para aquellas escuelas que atienden a los alumnos más pobres, considera distintos grados de autonomía en el uso de estos recursos adicionales, según el puntaje promedio alcanzado por la escuela en el SIMCE en relación al puntaje del propio grupo socio-económico. Adicionalmente, incorpora una cláusula que señala que una vez que se establezcan estándares estos se usarán en forma complementaria al puntaje promedio; es decir, no solo se considera la relación del puntaje promedio respecto del puntaje promedio del grupo socio-económico, sino también la distribución de alumnos de acuerdo a los estándares de desempeño. En esta ley se establece además que todos los centros educativos se clasificaran en tres tipos de establecimientos según sus resultados: autónomas, emergentes y en recuperación. En definitiva es una ley que intenta generar mecanismos de *responsabilización* y condicionar la autonomía del establecimiento según sus resultados en las pruebas nacionales.

4.4. Nuevo trato a la educación pública

El tercer pilar dice relación a cómo enfrentar el conjunto de falencias estructurales, que sin duda afectan la calidad y la equidad del sistema educativo municipal o de educación pública.

La constatación en este campo es que mientras más pobre y aislada la población escolar, es decir, mientras más necesitada del mejor servicio educativo, es más probable que su educación esté administrada con recursos humanos, materiales y financieros insuficientes. Hoy muchos municipios no tienen una estructura de ingresos que les permita dotarse de los recursos humanos suficientes para cumplir en forma óptima lo que se le pide a un sistema de administración escolar. Hasta el momento, no existen asociaciones de municipalidades que agreguen sus recursos para alcanzar una escala operacional que les permita gestionar eficiente y efectivamente escuelas pequeñas en zonas de relativo aislamiento y de amplia extensión geográfica.

En un número importante de comunas además, ve dificultado el financiamiento de sus establecimientos escolares por tres grandes causas: el pasaje progresivo de la matrícula al sector particular, que genera un inmediato efecto económico en los sostenedores municipales que reciben cada vez menos subvención estatal; el alto nivel de edad de los docentes municipales que, al estar regidos por el Estatuto Docente, aumentan sus salarios según la antigüedad y, la caída del crecimiento demográfico de la población en edad escolar, que comienza a sentirse y se espera sea aún más profunda en los próximos años. Se agrega además, que las escuelas municipales a diferencia de las subvencionadas particulares, no instalan mecanismos de selección del alumnado, lo que les implica recibir estudiantes de mayor vulnerabilidad y de menor rendimiento escolar, lo que demanda mayor gasto.

Así, la actual educación pública requiere un cambio significativo. En lo referido a Gobernabilidad y Responsabilización se deben redefinir funciones y aclarar responsabilidades; establecer estándares y metas, con consecuencias: incentivos y sanciones para las administraciones educacionales municipales o

públicas; y centralmente se deben entregar competencias técnico pedagógicas a estas administraciones (supervisión de establecimientos), lo que implica también redefinir funciones del MINEDUC. En el ámbito de las capacidades técnicas y gestión de las administraciones se requiere profesionalizar equipos, fortalecimiento de las capacidades técnico pedagógicas de ese mismo personal; así también la instalación de dispositivos para la identificación, irradiación y difusión de buenas prácticas (por ejemplo trabajo en redes). Para mejorar las actuales estructuras de costo se deben diseñar nuevos e implementar los actuales dispositivos que permitan alivianar las plantas docentes abultadas y adicionalmente, el desarrollo de nuevos y más efectivos mecanismos contables (por ejemplo la metodología de centro de costo).

Sean cual sean las opciones que se tomen en esta campo, resulta claro que el nuevo modelo Institucional para la Educación Pública, deberá cumplir ciertos "mínimos", tales como: a) capacidad para gestionar la entrega de apoyo técnico pedagógico a los establecimientos, integrándolo con su función administrativa; b) capacidad para asumir mayor autonomía (diseñar e implementar la política educativa, responsabilizándose por sus resultados) y c) lograr un cierto nivel de resultados educativos y financieros.

5. DESAFÍOS QUE SE ABREN CON LA NUEVA ARQUITECTURA DEL SISTEMA EDUCACIONAL

El primer desafío que este nuevo marco normativo presenta en el campo de los estándares, será cómo *asegurar la alineación y concordancia* entre el currículum y el sistema de evaluación de la calidad. El nuevo marco legal afecta de dos maneras el actual trabajo, por una parte, desvincula organizacionalmente la función evaluación de la de desarrollo curricular y formulación de estándares actualmente unidas organizacionalmente y, por la otra, agrega a la dimensión evaluativa un nuevo tipo de estándar, de desarrollo incipiente en Chile, el este marco legal denomina "*estándares indicativos de desempeño*", que se asemejan más a los que Ravitch (1996) denomina estándares de oportunidad para aprender.

Será así tarea central de la nueva institucionalidad la búsqueda de la coherencia entre estos instrumentos, todos potentes mensajes curriculares, ello exigirá a las distintas instituciones Ministerio y Agencia- un nivel de coordinación que en nuestro contexto muy pocas organizaciones públicas exhiben en la actualidad. Por otra parte, el desarrollo de una supervisión evaluativa, basada ahora en estándares indicativos de desempeño de las instituciones escolares, sus administradores, los docentes y directivos podría aportar -si existe coordinación eficiente- a un enriquecimiento y retroalimentación mucho más poderoso que el presente respecto, por ejemplo de, la implementación curricular, la cobertura curricular en distintos establecimientos, efectos del uso del currículum, etc.

Chile hoy camina hacia el desarrollo de componentes de su política en educación alineados: estándares y currículum nacional; estándares, evaluación de aula y medición nacional. El currículum chileno busca estar alineado con SIMCE, los textos escolares provistos por el Ministerio al sistema, la evaluación docente y estrategias de desarrollo profesional docente también poseen ese objetivo. Al mismo tiempo, al establecer un sistema de accountability o responsabilización de las escuelas y liceos, se crean nuevos modos de alineamiento, ahora con el sistema de medición de la calidad SIMCE, vinculados a incentivos tanto positivos como negativos (sanciones). Esta resulta una doble tensión a resolver en la implementación de la nueva institucionalidad.

Un segundo desafío, como ya se mencionó antes en el aparatado referido a los nudos, será conjugar adecuadamente las presiones o exigencias y los necesarios apoyos. El mejoramiento de la calidad, pero de calidad para todos, no se logrará sólo mediante la implementación de Mapas de Progreso de aprendizaje y de las mediciones SIMCE basadas en Niveles de Logro, para que ello ocurra es necesario que los estándares vayan acompañados de políticas que incidan en los factores condicionantes para una adecuada implementación. Lo mismo ocurrirá con los estándares indicativos de desempeño. El uso de los estándares sólo para control externo provoca una tensión permanente y generan resistencia para diferentes grupos de interés.

Entonces hay que hacerse cargo de la caricatura de los estándares como homogenizadores, los que al medir a todos con la misma vara, resultan inequitativos y discriminatorios, no ayudando a aumentar la igualdad de oportunidades. A igualdad de logros esperados, dicen los que cuestionan en esta materia, el esfuerzo que deben hacer los sectores más débiles para alcanzarlos es mucho mayor que el que deben hacer los bien dotados, con lo cual los estándares se convierten en una sobre-exigencia para los grupos más débiles, lo que a la larga los llevará a producir mayores fracasos, al facilitar su estigmatización negativa. El establecimiento de altos estándares desmotiva a los estudiantes de los grupos más vulnerables, confirmando lo que ya se sabe acerca de su modesto desempeño.

La tarea es lograr el adecuado equilibrio entre estándares exigentes y la posibilidad de que sean alcanzables. Es recomendable establecer estándares altos pero alcanzables y evitar así que el trabajo de los profesores se centre en mínimos que se vuelven máximos. Para tales efectos, considerar los actuales rendimientos como "línea de base" es un buen patrón de referencia.

Si existen exigencias desafiantes deben existir los apoyos proporcionales a esas exigencias. Ello es parte esencial de una política que instala estándares. Se requiere definir las instituciones que presten apoyo, desarrollar las capacidades de éstas y fomentar el liderazgo en cada institución educativa y a lo largo de todo el sistema. El tema no es el estándar, sino el camino escogido para avanzar hacia su logro y para eso se requiere mucha capacidad y liderazgo técnico pedagógico distribuido en todos los territorios y escuelas.

Un tercer desafío se refiere a los recursos financieros para estas políticas. Asegurar la calidad implica recursos y abundantes. La lógica tras este debate es que si el Estado exige resultados, en la forma de cumplimiento de estándares, debe entregar los recursos para que esto sea posible. La generación por parte del Estado de condiciones reales de posibilidad, introduce un tercer tipo de estándar que se une a los anteriores de desempeño y contenido, los estándares de oportunidad para aprender. Ello se traduce en que cada centro escolar posea los recursos adecuados para que los estudiantes sean capaces de satisfacer estándares de contenido y de desempeño desafiantes.

Una última reflexión, respecto del desafío de complejizar el concepto de calidad, el cual posee significado y contexto histórico y por tanto evoluciona en el tiempo. En este sentido el nuevo marco legal chileno combina como factores de la calidad a los estándares de aprendizaje y los estándares indicativos de calidad, que capturan también el proceso educacional, así no se trata solo de aprender, sino de aprender en buenas condiciones, en un entorno escolar democrático que permite el pleno despliegue de los talentos de todos los alumnos y todas las alumnas, asegurando una formación para vivir plenamente en la compleja y demandante sociedad del conocimiento. Al instalarse los estándares, se inicia nuevamente un ciclo, se vuelve a plantear la interrogante respecto de los factores que afectan la calidad,

reabriéndose ahora con evidencia el eterno debate sobre la calidad y dentro de ella de los estándares y los límites de estos.

REFERENCIAS BIBLIOGRÁFICAS

- Gysling, J. y Meckes, L. (2007). *Mapas de Progreso y Niveles de Logro SIMCE. Desarrollo de Estándares de Contenido y Desempeño en Chile: 2002 a 2007*. Documento de trabajo interno de la Unidad de Currículum y Evaluación del MINEDUC de Chile.
- Hargreaves, A., Earl, L., Moore, S. y Manning, S. (2001). *Aprender a cambiar. La enseñanza más allá de las materias y los niveles*. Barcelona: Ediciones Octaedro.
- Linn, R.L. y Herman, J.L. (1997). *La evaluación impulsada por estándares: Problemas técnicos y políticos en la medición del progreso de la escuela y los estudiantes*. Grupo de Trabajo sobre Estándares y Evaluación. PREAL.
- Ministerio de Educación (2003a). *Desempeños a Lograr en Lectura, Escritura y Matemática: 2º Nivel de Transición, Nivel Básico 1, Nivel Básico 2*, Campaña de Lectura, Escritura y Matemática.
- Ministerio de Educación (2003b). *Evaluación de Aprendizajes para una Educación de Calidad*, Comisión para el Desarrollo y Uso del Sistema de Medición de la Calidad de la Educación.
- OCDE (2004a). *Chile: Revisión de políticas nacionales de educación*. Chile: Ministerio de Educación.
- OCDE (2004b). *What Makes School Systems Perform: seeing school systems through the prism of PISA*. Paris: OCDE.
- OCDE (2005). *School factors related to quality and equity*. Paris: OCDE.
- PREAL (1999). *Desarrollo de Estándares nacionales y evaluaciones: tras la meta de mejor educación para todos*. Santiago de Chile: PREAL.
- Ravitch, Diane (1996). *Estándares Nacionales en Educación (Estado de la práctica)*. Santiago de Chile: PREAL.

LA EXPERIENCIA DEL SISTEMA NACIONAL DE EVALUACIÓN DEL PROCESO EDUCATIVO, SNEPE, EN PARAGUAY. APRENDIZAJES Y DESAFÍOS

Marta Lafuente

Revista Iberoamericana de Evaluación Educativa 2009 - Volumen 2, Número 1

<http://www.rinace.net/riee/numeros/vol2-num1/art3.pdf>

Fecha de recepción: 15 de marzo de 2009

Fecha de comunicación de dictamen: 18 de marzo de 2009

Fecha de aceptación: 23 de marzo de 2009

1. CONTEXTO NACIONAL Y REFORMA EDUCATIVA

El Paraguay según la DGEEC¹ (2003) es un país multilingüe y pluricultural de 6.054.976 habitantes. Cuenta con una superficie es de 406.752Km². Está situado en el hemisferio sur del continente americano, entre los paralelos 19° 18' y 27° 3' de latitud sur y los meridianos 54°15' y 62° 38' de longitud, al oeste de Greenwich. Limita al norte con Bolivia y Brasil, al este con Brasil y Argentina, al sur con Argentina y al oeste con Bolivia.

Existe una alta concentración de habitantes en la Región Oriental que dispone de un 40% del territorio con un 97,4% de la población, mientras la Occidental o Chaco Paraguayo, concentra el 60% del territorio restante con una población del 2,6%.

El idioma guaraní se habla en el 59,2% de los hogares mientras el castellano se habla en el 35,7%. En las ciudades, el castellano es hablado por el 54,7% de los casos mientras que en el campo la lengua más hablada es el guaraní, alcanzando el 82,7% de los hogares como lengua predominante.

La peculiaridad más resaltante del comportamiento lingüístico es que el guaraní, siendo lengua indígena, es utilizado como instrumento de comunicación verbal por la mayoría de la población no indígena. Esta característica le diferencia de las demás naciones donde se conservan idiomas amerindios. El 2% de la población es indígena, distribuidos en 20 pueblos y cinco familias lingüísticas.

La población 0 a 19 años en el 2006 representaba el 46% de la población total del país, lo cual indica una fuerte demanda para el sector educativo, sin embargo, el Censo Nacional de Población y Viviendas 2002 muestra una desaceleración del ritmo de crecimiento poblacional, lo cual denota una reducción de la demanda futura por la educación.

De acuerdo al SERCE (2008) el índice de desigualdad en la distribución del ingreso del Paraguay, medido por el Coeficiente Gini, es de 0.54 y su Índice de Desarrollo Humano 0.755. Estos indicadores dan cuenta de los niveles de desigualdad y pobreza del país.

Creemos importante contextualizar el proceso vivido en el país en la década de los 90 para hacerse una idea más clara de las características de la reforma Educativa. Según Rivarola (2000) el Paraguay en el año 1989 se encontró ante una situación histórica imprevista y singular con el abrupto desmoronamiento del régimen totalitario que imperó 35 años por cuanto se inaugura cambios en el funcionamiento del Estado, el manejo del aparato gubernamental, las formas y niveles de participación ciudadana en las cuestiones públicas y el carácter de las instituciones. Refiere el autor que la ciudadanía por primera vez toma conciencia de estar viviendo una etapa excepcional y por la decidida voluntad de gestar profundas transformaciones en la sociedad paraguaya. El país con una larga tradición autoritaria y sin instituciones democráticas tuvo que enfrentar una transición política, económica, social, cultural compleja y muy diferente de otros países de la región que también vivieron largos periodos de dictadura.

De acuerdo a este autor se destacan dos aspectos centrales de la vida en libertad en el país, la profunda fe en la democracia y la alta expectativa en la educación como la estrategia para afianzar la democracia y la posibilidad de construir una nueva sociedad.

¹ DGEEC Dirección General de Estadísticas Encuestas y Censos

La Reforma educativa surge como la primera voz en libertad del pueblo paraguayo ante la caída de la dictadura de Stroessner. Podría decirse que ante la disyuntiva de la continuidad o el cambio, la elección de emprender el proceso de profunda transformación en el sistema educativo fue un imperativo para los paraguayos. El carácter participativo de la reforma quizá pueda identificarse como el gran deseo de protagonismo de la gente en un nuevo estilo de vida social y política: la democracia. El primer informe de avance del Consejo Asesor de la Reforma Educativa (1992) se titula "Reforma Educativa: Compromiso de todos", lo cual muestra la significación de la participación en el proceso. Este sentido se expresó en aquellos tiempos muy elocuentemente en guaraní "Ñandeha" pronombre incluyente como expresión de protagonismo activo en la participación y quizá su uso estaba inaugurando la visión de espacio compartido de lo público, aquello que es de todos. Por demasiados años se vivió la idea de los unos contra los otros sin la experiencia de lo compartido tan intrínseco a la educación como bien público.

Los diagnósticos mostraban claramente los efectos de la educación bancaria² sustentada en la transmisión vertical de saberes con escasa pertinencia cultural para el perfil sociolingüístico del país, la carencia de sistemas de información científica y de la escasa cultura de la evaluación. En el modelo de la educación bancaria se privilegió la transmisión de "qué" antes que el conocimiento del "cómo". Se enfatizó el aprendizaje memorístico, descuidando el aprendizaje comprensivo y más aún el aprendizaje reflexivo y crítico. Se caracterizó por un fuerte control ideológico del sistema educativo y según Rivarola (2000), la educación paraguaya carecía de una mínima cultura crítica para iniciar su transformación.

Con la democracia en el Paraguay la educación adquirió gran relevancia y esto se plasma en la Constitución Nacional del año 1992. Por primera vez una constitución de la República declara al guaraní y al castellano como lenguas oficiales y garantiza la enseñanza en lengua materna.

Hasta el año 1993 el Ministerio de Educación experimentó cambios frecuentes de autoridades en un clima de cultura autoritaria y reacia a la nueva demanda ciudadana de profunda transformación de la sociedad. Se podría decir que estos primeros años transcurrieron en medio de tensiones y conflictos entre las altas expectativas de cambio de la ciudadanía y la resistencia interna del aparato de control formado en la doctrina de la seguridad nacional.

La ciudadanía exigía la reforma del sistema y demandaba en la conducción la presencia de personalidades con legitimidad por su pensamiento democrático y su solvencia profesional. En ese contexto se conformó el Consejo Asesor de la Reforma Educativa (CARE) integrado por doce miembros.

La cooperación internacional también ofrecía apoyo técnico y financiero al país; en el 94 se firmó un convenio de préstamo con el Banco Interamericano de Desarrollo para la Educación Básica y posteriormente con el Banco Mundial para el financiamiento la educación media.

En el año 1995 el Ministerio de Educación y Cultura (MEC), firmó un acuerdo de asistencia técnica con el Instituto de Desarrollo Internacional de la Universidad de Harvard con el objeto de fortalecer la capacidad de uso de la información para la formulación de políticas. El Centro de Estudios Sociológicos (CPES) y el HIID prepararon un estudio sobre la situación de la educación como base para la formulación de una propuesta de la nueva educación paraguaya.

² En el sentido del pedagogo brasileño Paulo Freire.

Según Reimers (2000) este acuerdo permitió fortalecer cuatro instancias del MEC: el Grupo de Diálogo Estratégico, la Unidad de Análisis de Política, la de Planeación y una Unidad de Investigación. Como resultado de este proceso se elaboró el Plan Estratégico Paraguay 2020, el mismo se constituyó en un instrumento de gestión muy importante para la reforma.

En 1998, se promulga La Ley N° 1264 General de Educación, que establece los principios generales para la educación pública y privada del Paraguay y otorga a la educación un lugar prioritario para la consolidación de la democracia, la disminución de desigualdades y la apertura de nuevas oportunidades para todos los habitantes del país. Dicha ley recoge el espíritu de los debates de esos años sobre la responsabilidad del estado de contar con mecanismos de evaluación para la educación y la necesidad de informar a la sociedad en forma permanente sobre los procesos y resultados.

1.1. Propósitos de la Reforma y concepción educativa

Entre los desafíos asumidos por la Reforma para impulsar la modernización del país de manera de lograr una inserción exitosa en la dinámica de las sociedades progresistas de la región y del mundo está el contribuir para la desaparición de los elementos que no hacen sino obstaculizar el avance hacia el futuro y a la construcción de conocimientos y capacidades.

La Reforma se inspira en la Declaración y el Marco de acción aprobados por la Conferencia Mundial sobre Educación para Todos³ como una importante fuente de referencia en cuanto a la nueva concepción de educación entendida como satisfacción de las necesidades básicas de aprendizaje de los niños, jóvenes y adultos como la base para el desarrollo humano permanente.

Se observa que la Reforma desde esta nueva concepción para el desarrollo de las personas orientada a formar competencias básicas e incentivar el aprendizaje permanente pretende superar la concepción educativa de puras transmisiones de “lecciones” que no preparan para la vida sino para los exámenes y la desocupación o en mejor de los casos para ocupaciones muy específicas de poca transferencia, flexibilidad y adaptación.

Las principales orientaciones de las políticas educativas ⁴(MEC.CARE 1998) que orientaron las prioridades de la Reforma en la década del 90 han sido las siguientes:

- Universalizar el acceso a la educación y fomentar la equidad
- Priorizar la calidad de la educación
- Atender la pertinencia curricular, (educación bilingüe) mejorar la eficiencia y la eficacia del sistema
- Ampliar el alcance y los medios educativos
- Formar y capacitar los recursos humanos
- Ampliar y mejorar la infraestructura
- Implementar programas focalizados para la discriminación positiva de escuelas de alto riesgo.

³ Declaración Mundial sobre Educación para Todos. Jomtien, Tailandia. 1990.

1.2. Breve diagnóstico de la situación educativa

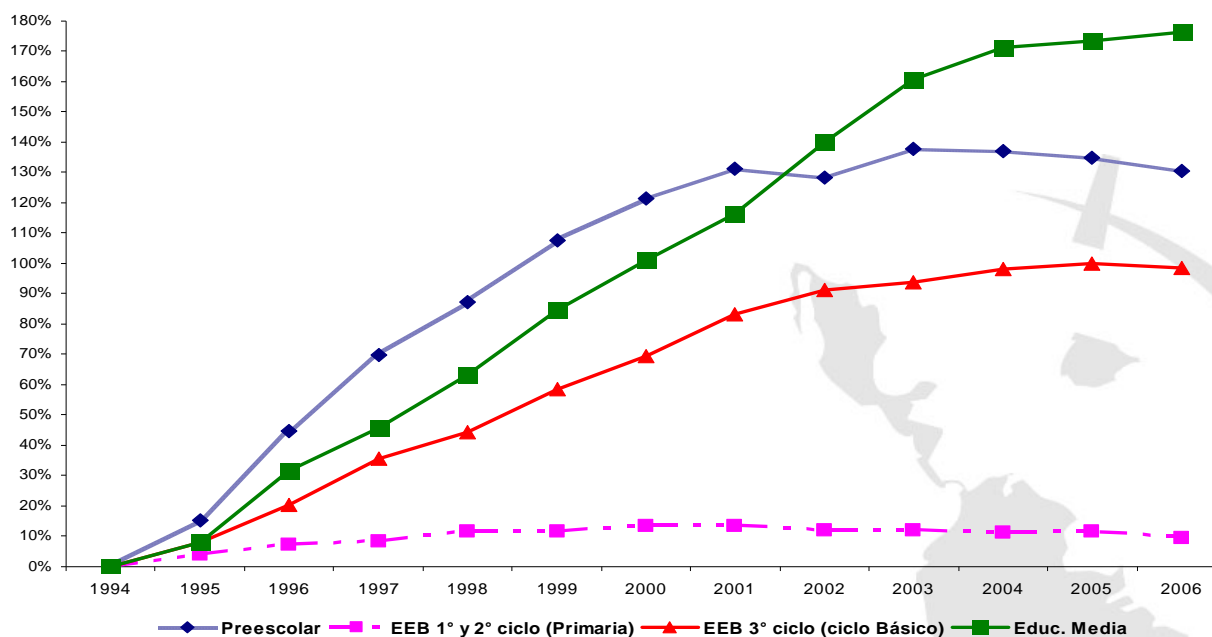
El cuadro muestra en números absolutos y relativos la expansión del sistema entre los años 1994 y 2006 en todos los ciclos; principalmente en el pre escolar, el tercer ciclo y la educación media. El primero y segundo ciclos por ser el tramo obligatorio a lo largo de la historia del sistema educativo presenta un comportamiento más estable. En el año 1994 se inicia la reforma educativa en el aula con nuevo currículo y textos gratuitos en las escuelas de gestión oficial y subvencionada.

CUADRO 1. MATRÍCULA POR NIVEL EDUCATIVO. AÑO 1994-2006

Año	Preescolar	EEB 1° y 2° ciclo (Primaria)	EEB 3° ciclo (ciclo Básico)	Edu. Media	Total
1994	51.671	835.089	159.675	76.239	1.122.674
2006	119.128	913.056	317.130	210.623	1.559.937
Crec. Absoluto 2003-2006	67.457	77.967	157.455	134.384	437.263
Crec. Porcentual 2003-2006	131%	9%	99%	176%	39%

Fuente: MEC, DGPEC, SIEC 1994-2006.

GRÁFICO 1: INCREMENTO PORCENTUAL DE LA MATRÍCULA POR NIVEL EDUCATIVO PERIODO 1994-2006. AÑO BASE 1994



FUENTE: MEC, DGPEC, SIEC 1994-2006.

El cuadro 2 muestra el impacto de la pobreza en el acceso a la educación. Si bien existe una mayor oportunidad se observa que los estudiantes que residen en zonas urbanas acceden en mayor proporción que los niños y niñas de zonas rurales. En este contexto los que tienen menor acceso son los niños y niñas que viven en zonas rurales y proceden de hogares pertenecientes a los quintiles de menores ingresos.

CUADRO 2. POBLACIÓN DE 5 A 18 AÑOS POR QUINTIL DE INGRESO, SEGÚN ASISTENCIA A INSTITUCIÓN EDUCATIVA DE ENSEÑANZA FORMAL

Área de residencia	Condición de asistencia	Quintiles de ingreso					Total de grupo
		I	II	III	IV	V	
Urbana	Si asiste	82%	88%	91%	91%	94%	89%
	No asiste	18%	12%	9%	9%	9%	11%
	Total	100%	100%	100%	100%	100%	100%
Rural	Si asiste	73%	78%	81%	85%	82%	78%
	No asiste	27%	22%	19%	15%	18%	22%
	Total	100%	100%	100%	100%	100%	100%
Total	Si asiste	76%	83%	87%	88%	91%	84%
	No asiste	24%	17%	13%	12%	9%	16%
	Total	100%	100%	100%	100%	100%	100%

FUENTE: DGEEC, EPH 2005, Elaboración del MEC

Desde el año 94 la edad de ingreso al pre escolar es de 5 años y al primer grado 6 años. El siguiente cuadro nos muestra un importante cambio en las tasas netas del pre escolar, el tercer ciclo y un poco menos en la educación media. Aún con la evolución se observa que por razones culturales y también de accesibilidad se retarda la llegada a los primeros años.

CUADRO 3. TASA BRUTA Y NETA DE ESCOLARIZACIÓN POR NIVEL EDUCATIVO. PERIODO 1994-2006

Año	Preescolar		Educación Escolar Básica				Educación Media	
			1° y 2° Ciclo		3° Ciclo		Ciclo Bachillerato	
	Tasa bruta	Tasa neta	Tasa bruta	Tasa neta	Tasa bruta	Tasa neta	Tasa bruta	Tasa neta
1994	38	30	112	92	50	35	28	20
2006	83	67	108	92	78	56	53	39

FUENTE: MEC, DGPEC, SIEC 1994-2006

El siguiente cuadro evidencia la brecha existente en el promedio de años de estudios entre las zonas urbana y rural en diferentes tramos de edad, si bien la reforma ha impactado positivamente en el rango de menores de 25 años.

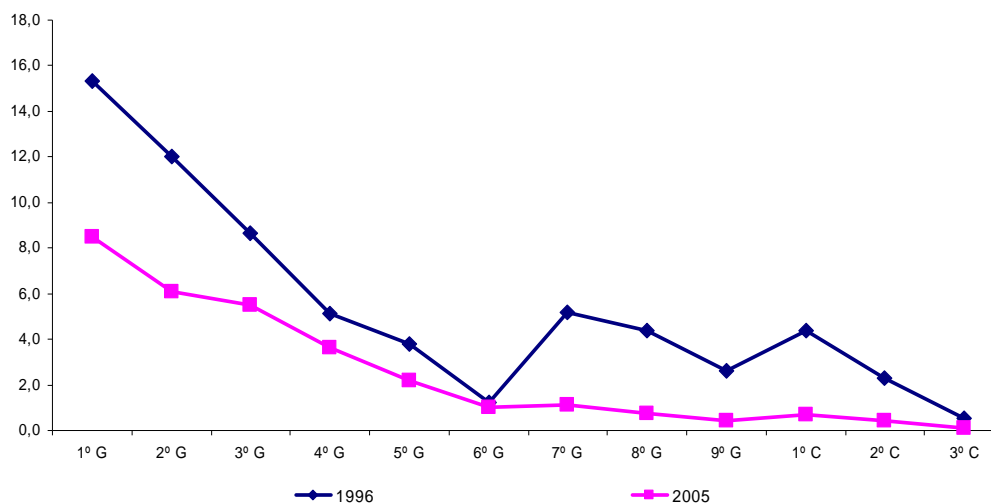
CUADRO 4. PROMEDIO DE AÑOS DE ESTUDIO POR ZONA. AÑO 2005

Área de residencia	Edad			Total
	25 años	45 años	65 años	
Urbana	10.5	8.3	7.8	9.4
Rural	7.2	4.9	2.7	5.5
Total	9.4	6.8	5.6	7.9

FUENTE: DGEEC, EPH 2005, Elaboración del MEC

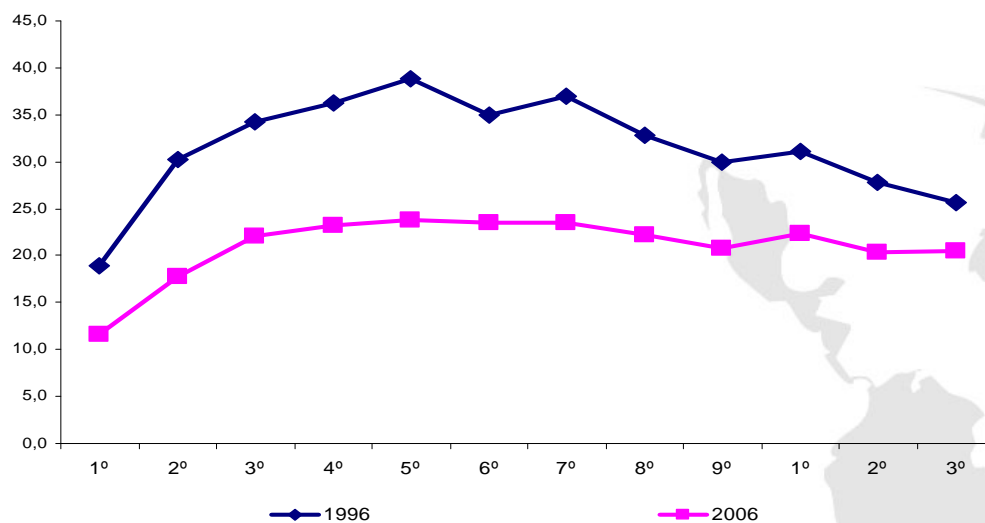
- Persiste la repitencia en los dos primeros ciclos de la EEB a pesar de su disminución. Los gráficos 2 y 3 muestran el comportamiento de los indicadores de ineficiencia del sistema. La sobriedad se debe más a la repitencia que al acceso tardío a la escuela.

GRÁFICO 2. TASA DE REPITENCIA POR GRADO. AÑO 1996-2005*



* Se dispone de datos de repitencia a partir de año 1996
FUENTE: MEC, DGPEC, SIEC 1994-2006

GRÁFICO 3. TASA DE SOBREEDAD POR GRADO. AÑO 1996-2006

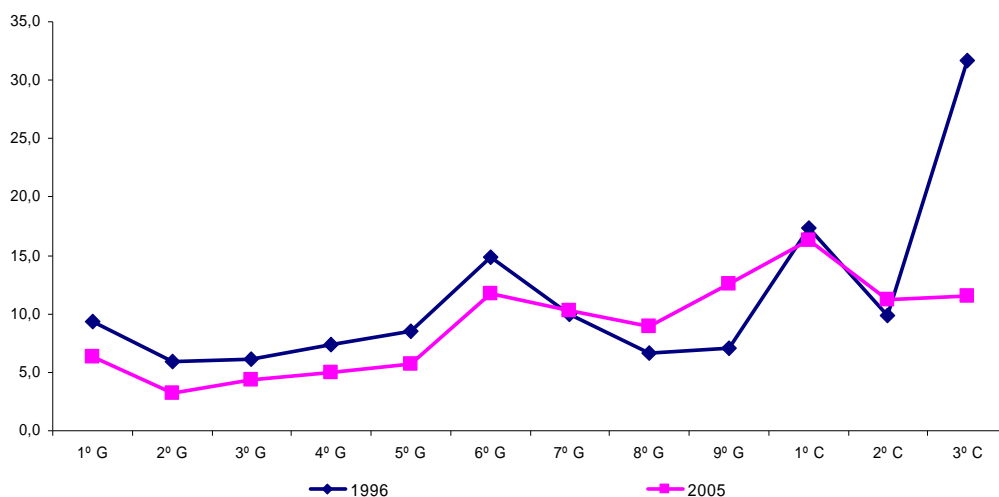


FUENTE: MEC, DGPEC, SIEC 1994-2006

- Persiste la deserción
- El sistema en general muestra una mayor capacidad de retención en cada Cohorte (1990/2001: 27%); la misma tendencia se observa en cuanto a rendimiento (1990/2001: 20%). Esto es, de

cada 100 alumnos que se inscribieron en el primer grado en el año 1990 27 llegaron al último año de la educación media y 20 egresaron en noviembre del año 2001.

GRÁFICO 4. TASA DE DESERCIÓN POR GRADO. AÑO 1996-2005

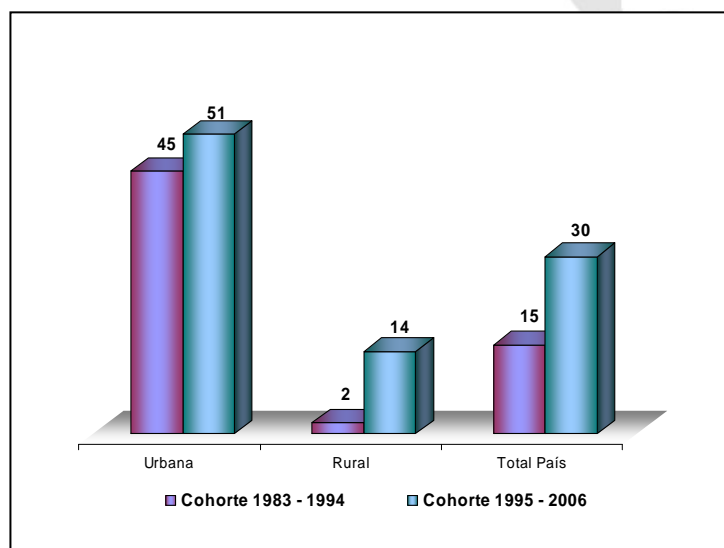


FUENTE: MEC, DGPEC, SIEC 1994-2006

Los gráficos 4 y 5 dejan observar el incremento de la retención en los doce años de escolarización (1º grado de EEB hasta 3º Curso de la Educación Media). En Cohortes 1983-1994 y 1995-2006, la cantidad de personas que consiguen alcanzar 12 años de escolarización se ha duplicado a nivel nacional, mientras que en zonas rurales se ha incrementado 7 veces más, el crecimiento más bajo se registra en zonas urbanas.

Según la Dirección de Planificación Educativa el mejoramiento de los indicadores de eficiencia no ha sido de la magnitud esperada para lograr el cumplimiento de las metas establecidas en el Plan Paraguay 2020.

GRÁFICO 5. TASA DE RETENCIÓN POR ZONA Y COHORTE



FUENTE: MEC, DGPEC, SIEC 1996-2005. Se refiere a cohorte de 12 años de escolaridad

2. EL SISTEMA NACIONAL DEL PROCESO EDUCATIVO, SNEPE

2.1. El marco institucional

El Sistema Nacional de Evaluación del Proceso Educativo (SNEPE) es creado en abril del año 1995 por Resolución Ministerial⁵, se instala en la Dirección de Orientación que a partir de esta institucionalización pasa a denominarse Dirección de Orientación y Evaluación Educativa. El propósito es generar información permanente, válida, confiable y oportuna sobre el nivel de logro de los aprendizajes alcanzado por los alumnos y alumnas, así como sobre las variables complementarias o contextuales que ayudan a interpretar mejor los resultados de las mediciones del rendimiento académico.

Los proyectos de mejoramiento de la educación básica y educación media financiados con créditos del BID y del Banco Mundial respectivamente tenían como componentes la medición del rendimiento académico. Estos organismos incorporaron en la agenda de la educación paraguaya la evaluación de la calidad y la rendición de cuentas públicas de los resultados. Se promovieron debates nacionales con expertos internacionales, académicos locales y autoridades sobre los distintos aspectos de la evaluación: el marco institucional, las ventajas y desventajas de situarlo dentro o fuera del Ministerio de Educación, el modelo de pruebas, los requerimientos técnicos y tecnológicos, etc.

Como resultado de esos debates se tomó la decisión de diseñar e instalar el nuevo sistema dentro del MEC. De esa manera se genera el marco institucional desde donde se gestionan los componentes de los proyectos del BID y del Banco Mundial en un único sistema de evaluación dado que los demás componentes eran gestionados desde unidades coordinadoras fuera del Ministerio. El diseño del SNEPE contemplaba una proyección al 2002.

En esos años en América Latina sólo Bolivia y Paraguay carecían de sistemas de evaluación. El SIMCE de Chile era el principal referente regional en materia de evaluación de la calidad. Por ese tiempo se fundaba el Laboratorio Latinoamericano de Evaluación de la Calidad de la Educación, LLECE.

El LLECE fue un espacio técnico político muy importante para el desarrollo del sistema paraguayo por ser una instancia de intercambio y desarrollo profesional, de discusión de la responsabilidad de la evaluación y fundamentalmente se constituyó en una instancia de diálogo y construcción de un marco conceptual regional. Ese contexto sirvió de apoyo al Paraguay en un sistema educativo sin cultura ni tradición en evaluación. Hay que tener en cuenta que la principal política de los 90 ha sido la provisión de insumos para las escuelas, textos, capacitaciones, nuevos programas de estudios.

El grupo de diálogo estratégico instaló la discusión sobre la toma de decisiones informada lo cual generó la necesidad de construir un sistema de información con dos fuentes: las estadísticas continuas y el rendimiento académico.

En el año 2003 coincidente con un nuevo de periodo de gestión del MEC y como resultado de evaluaciones institucionales se decide separar al SNEPE de la Dirección de Orientación y pasa a ser una nueva Dirección

⁵ Resolución N° 687 del 6 de abril de 1995. En Propuesta para el sistema nacional del proceso educativo. SNEPE Ministerio de Educación y Cultura. Rca del Paraguay.

dependiente de la Dirección General de Desarrollo Educativo al mismo nivel que Currículum y Orientación. Esto le ha permitido mayor autonomía en la gestión.

2.2. Cobertura, áreas y grados y enfoque metodológico

En los inicios se optó por una cobertura muestral de la evaluación considerando los propósitos, la viabilidad financiera y también para disminuir posibles resistencias de los docentes. La primera evaluación se llevó a cabo en octubre del año 1996⁶. Por muchos años se evaluaron los finales de ciclo de la Educación Escolar Básica 3°, 6° y 9° grados y el tercer curso, último año, de la Educación Media.

El Snepe se ha iniciado con el uso de la teoría clásica (TCT) y desde el 2004 a incorporado la Teoría de Respuesta al Ítem (TRI). El cambio del enfoque se dio por los avances de la evaluación, la formación de los técnicos nacionales y sobre todo por las expectativas de mejor aprovechamiento de los resultados para la orientación del aprendizaje. De esta manera se corresponden las orientaciones del currículo nacional por competencias y la evaluación del aprendizaje.

En la educación media se dispone de publicaciones con el modelo TRI (MEC 2007) sin embargo en la Educación Escolar Básica se tienen validados los instrumentos y en la próxima evaluación nacional debiera adoptarse como modelo de este nivel también.

Para la evaluación del rendimiento se utilizan pruebas de opción múltiple construidas en base al currículo nacional, pruebas de redacción y cuestionarios complementarios para estudiantes, padres y madres, docentes y directores.

En la educación inicial se adaptaron instrumentos internacionales para estimar el desarrollo del niño y evaluar el contexto. Para la evaluación del pensamiento crítico se elaboró un test que mide procesos deductivos, procesos inductivos y proceso crítico-sociales.

Todos los instrumentos son sometidos a estudios de validación cuyos indicadores son publicados en los informes de resultados. Con este cambio se pretende mejorar la calidad de la información para la toma de decisiones y hacer comparables las sucesivas evaluaciones. En el futuro se podrían realizar estudios de cohortes.

Desde su creación el sistema de evaluación tiene fin diagnóstico, de orientación de políticas y de verificación del impacto de programas educativos a nivel nacional.

2.2.1. En la Educación Escolar Básica

Se evalúan Comunicación, Matemática cada dos años, Vida Social y Trabajo y Medio Natural y Salud en forma alternada. Las evaluaciones del rendimiento académico se acompañan de cuestionarios de actitudes de los niños, cuestionarios para directores, docentes, madres y padres de familia.⁷

El currículo de la Educación escolar Básica ya era resultado de la reforma del nivel por tanto era relativamente sencillo elaborar las tablas de especificaciones con las competencias de cada ciclo. El diseño curricular orientaba adecuadamente lo que se debía evaluar.

Las pruebas se construyen sobre la base de un análisis sistemático de los contenidos centrales en cada una de las áreas evaluadas; la aplicación de dichas pruebas proporciona información sobre el nivel de logros

⁶ MEC(1997) Primer informe de Resultados. SNEPE.

⁷En la bibliografía se citan todos los informes de resultados. Ver también en www.arandurape.edu.py

obtenidos por los estudiantes en las diferentes áreas. Hasta el año 2004 el modelo de análisis utilizado para estimar el rasgo de desempeño de los estudiantes era la Teoría Clásica del Test (TCT). A partir del año 2005 el SNEPE incorpora al modelo utilizado, la Teoría de Respuesta al Ítem (TRI) como complemento importante de la TCT en el diseño, construcción y evaluación de las pruebas. En la educación media todos los instrumentos fueron diseñados y procesados bajo dicho modelo (MEC 2007), en la Educación Básica los instrumentos han pasado por la prueba experimental.

2.2.2. En la Educación Media

La primera evaluación en este nivel del sistema se aplicó en el último curso de la educación media en el año 1988. Posteriormente al tercer y sexto cursos en el año 2000 con el propósito de contar con un diagnóstico para iniciar la reforma de la educación media. En 1998 y 2000 se evaluaron las áreas de lengua castellana, matemática y estudios sociales del plan común. El currículo de la secundaria tenía al menos 30 años de implementación y se tuvo que hacer un enorme y dificultoso esfuerzo para construir las tablas de especificaciones (MEC 1999) debido a la gran cantidad de cambios que han experimentado los programas de estudios a lo largo de tres décadas. Al punto de no contar prácticamente con un referente del currículo formal vigente ya que había tantos programas que se había perdido la noción de plan común. Se tuvo que aumentar y ampliar el trabajo de campo para cotejar el currículo en el aula con lo prescrito en los varios programas "oficiales". Esto ya daba un primer dato importante para el futuro de la reforma de la educación.

La evaluación de la secundaria tropezó con todas las resistencias por parte de los actores, era muy difícil acceder a informaciones en los colegios. La noción de lo público y de la rendición de cuentas no eran los rasgos que identificaban a las instituciones. Al menos dos factores incidían fuertemente, el hecho de ser un nivel muy selectivo y la falta de supervisión pedagógica del Ministerio. La supervisión en la dictadura tenía más un carácter de control ideológico que educativo. Según la tasa neta del año 2000, 31 de cada 100 estudiantes cursaban el último ciclo de la educación media⁸.

Fue muy interesante cotejar las actitudes de los educadores frente a la evaluación externa a la escuela. Mientras a los maestros de la Educación Escolar Básica les costaba aceptar que debían abandonar sus salas de clase para dejar entrar al examinador y sin su presencia en la Educación Media el rechazo era por la evaluación como tal. Costó trabajo sensibilizar y crear una cultura de apertura. No fue fácil instalar la idea que la evaluación serviría de gran apoyo para reorientar las prácticas en el aula.

2.3. Informes y difusión de los primeros resultados

Desde los inicios se elaboran dos tipos de informes, uno con informaciones de los resultados nacionales agregados y clasificados según variables consideradas en la estratificación de la muestra, tipos de gestión y turnos de los cursos evaluados, entre otras, dirigido a tomadores de decisiones, medios de comunicación y público en general. Este informe además contiene todo los datos de los instrumentos y procedimientos de análisis y un glosario⁹ No hay que olvidar que uno de los propósitos permanentes es el mejoramiento de la cultura de la evaluación. El otro informe tiene como destinatario a cada institución educativa participante de la muestra. El contenido incluye el resultado de la escuela en cada área evaluada y con referencia a los resultados de escuelas del mismo estrato, del mismo departamento geográfico y la ubicación de la

⁸ MEC (2007) Diagnóstico de la Educación en Paraguay 1994-2006 Dirección General de Planificación Educativa. Documento de trabajo.

⁹ En la bibliografía se citan todos los informes de resultados del SNEPE desde el año 1997.

institución respecto al rendimiento nacional. Ejemplos de estos informes se publican como anexo de los informes agregados.

El informe nacional generalmente se difunde en conferencias de prensa y con presencia de autoridades educativas del nivel central, supervisores de los diferentes departamentos geográficos y directores de instituciones. Cada año y según las prioridades de la agenda se organizan eventos especiales para profundizar la difusión o se difunde en seminarios, congreso u otras modalidades de reuniones de actores de la gestión educativa. Los informes nacionales de las últimas evaluaciones están disponibles en el portal de MEC.

Los informes de las instituciones se entregan a los directores en talleres en cada departamento geográfico con involucramiento de los supervisores. Posteriormente cada zona de supervisión se hace responsable del uso en las informaciones en talleres de capacitación docente. El nivel central del MEC no ha instalado un sistema de seguimiento del uso de dichas informaciones a nivel institucional por tanto se hace difícil dar un reporte al respecto.

2.4. Avances en los niveles, enfoques y áreas de la evaluación

El desarrollo de la agenda de transformaciones de la reforma supuso nuevos desafíos para el SNEPE en cuanto a (i) generar informaciones sobre otros niveles del sistema educativo como la formación docente y la educación inicial, (ii) y (ii) la incorporación de nuevas áreas y temáticas en la educación media.

Como se señalara anteriormente en los 90 las políticas estaban más centradas en la distribución de insumos por la gran expansión que experimentaba el sistema educativo. Sin embargo los programas de inicios del nuevo siglo plantearon con mayor énfasis la necesidad de monitorear los procesos y resultados de las intervenciones. Las informaciones estadísticas y las primeras evaluaciones del SNEPE mostraban los insuficientes resultados lo cual reforzaban dos ideas. La necesidad de focalización de las políticas para atender la equidad y otras dos relacionadas con la calidad del propio sistema de evaluación. Se lo veía muy concentrado en las mediciones de la Educación Básica y el modelo teórico de las pruebas ya no satisfacía las preguntas sobre la calidad, por tanto el requerimiento era de avance hacia otros enfoques con mayores potencialidades de orientación pedagógica. Esto ha llevado Por otra parte la formación docente era fuertemente cuestionada por la sociedad fundamentalmente los medios de prensa en situaciones de huelga docente en busca de reivindicaciones salariales o en el momento de la difusión de resultados del SNEPE Eso ha llevado a la gestión a profundizar en la complejidad de los cambios con nuevos programas para la formación inicial de educadores, la implementación de políticas y programas de educación inicial y la reforma la educación media.

2.4.1. Evaluación de la Formación inicial de Docentes

La Formación Docente en Paraguay es una carrera de nivel superior no universitario de tres años de duración, como requisito se requiere el título de educación media y aprobar exámenes de admisión. La formación se imparte en Institutos de Formación Docente de gestión oficial y de gestión privada. Con la expansión de la educación obligatoria este nivel educativo experimentó una masiva apertura de nuevos institutos y alta matrícula debida a la demanda de docentes. Tanto fue así que desde el año 2000¹⁰ se tuvo que suspender la apertura de nuevos institutos por la saturación de maestros titulados y desocupados, principalmente en la EEB en los dos primeros ciclos, primero a sexto grados.

¹⁰ MEC. Resolución 10747 del año 2000.

Luego de esta masificación hubo que intensificar las políticas de evaluación de la calidad de la formación inicial de los docentes y la implementación de programas de formación en servicio articulados con la formación inicial que hasta ese entonces operaban de manera separada. Los Institutos sólo se ocupaban de la formación inicial, mientras que la formación en servicio era responsabilidad del Ministerio de Educación a través de los programas de la Dirección de Formación Docente y de los programas de mejoramiento de la calidad con apoyos del BID y del Banco Mundial principalmente.

En este contexto se diseña y desarrolla una línea de acción de evaluación y mejoramiento de los Institutos de formación docente con tres componentes: evaluación censal del rendimiento académico de los estudiantes del último año de la carrera de los institutos de gestión oficial y privada, auto evaluación y proyectos concursables de mejoramiento de los institutos de gestión oficial.¹¹

La agenda presentaba dos importantes novedades: la evaluación de los institutos de gestión privada, no sin resistencias por ser la primera experiencia de un esfuerzo del estado por el control de calidad y, la evaluación censal convirtiéndose en el único nivel en el que se implementó esta cobertura. Se implementaron tres operativos, años 2001, 2003 y 2005.¹²

Las áreas evaluadas fueron Lengua Castellana, Lengua Guaraní, Producción de Textos, Matemática, Estudios Sociales, Ciencias y Tecnología.

CUADRO 5. PROMEDIO DE LOGROS NACIONALES 3º CURSO ESTUDIANTES DE FORMACIÓN INICIAL DE LOS IFD

Estrato	Lengua Castellana (%)	Lengua Guaraní (%)	Matemática (%)	Ciencias y Tecnología (%)	Estudios Sociales (%)
Oficial	44.24	52.4	44.95	51.54	51.39
Privado	42.66	50.34	44.23	52.67	51.22
Resultado Nacional	44.46	50.33	44.59	52.12	51.30

FUENTE: MEC, DGDE, Informe de Resultados SNEPE - IFD 2003-2004

CUADRO 6. LOGROS NACIONALES 3º CURSO IFD 2005

Porcentaje de respuestas correctas	
Área	%
Lengua Castellana	50.27
Lengua Guaraní	54.88
Matemática	50.63
Ciencias y Tecnología	60.15
Estudios Sociales	59.05

FUENTE: MEC, DGDE, SNEPE 2005/2006

Tal como se puede apreciar, los resultados mostraron las debilidades en la formación inicial tanto en las instituciones de gestión oficial como privadas en los conocimientos específicos de las áreas y en la didáctica de las disciplinas. Dichos resultados fueron difundidos cada año en informes nacionales de distribución masiva y a todos los institutos en encuentros interactivos y de reflexión sobre las implicancias para cada área del conocimiento en la formación inicial¹³.

¹¹ MEC (2001). Un horizonte nuevo para los Institutos de Formación docente. Asunción. Paraguay; MEC (2003) Manual de procedimientos para la Autoevaluación de los Institutos de Formación Docente. Asunción Paraguay.

¹² MEC (2002), (2004) y (2008). Informe de Resultados. Cómo redactan los futuros maestros. Asunción Paraguay

¹³ Idem 12

Las constataciones ayudaron a desarrollar políticas de formación continua mediante la articulación de todos los programas de la formación inicial y la formación en servicio. Al mismo tiempo, la formación en servicio se asumió como la nueva agenda en los Institutos de gestión oficial. Cabe señalar que hasta ese momento, la educación de los docentes una vez egresados de los Institutos se hallaba disperso en todo el sistema.

Para concretar estas políticas hubo que generar las condiciones en los Institutos a través de programas de mejoramiento, los que incluyeron dotación de recursos bibliográficos y tecnológicos, formación de formadores y diseño de programas de formación en servicio.

La Formación Docente Continua en Servicio comprende la: Profesionalización, Especialización, Habilitación Pedagógica para egresados universitarios, Capacitación, Actualización y Perfeccionamiento.

Todo este proceso ha sustentado y evidenciado la necesidad de pensar en el mejoramiento permanente de los Institutos por lo cual se avanzó en el diseño y validación del sistema de Licenciamiento y Acreditación de las instituciones formadoras de educadores. El Licenciamiento debe aplicarse desde el 2009 en forma experimental en un grupo de Institutos. El mecanismo de Licenciamiento estará a cargo del MEC con el acompañamiento de la Agencia Nacional de Acreditación, los resultados serán utilizados con fines formativos para el mejoramiento de las instituciones que forman maestros y como fase preparatoria para la Acreditación. La Acreditación será una fase posterior y a cargo de la Agencia Nacional de Evaluación y Acreditación de la Educación Superior, ANEAES¹⁴. El objetivo es evaluar la calidad de las instituciones formadora de educadores de forma a garantizar el cumplimiento de los criterios esenciales de calidad¹⁵.

La otra línea de transformación importante es la Evaluación del Desempeño para la Certificación Profesional de los educadores. Dicho mecanismo es definido como *el conjunto de procesos que acompaña el desarrollo profesional del educador, asegurando que el mismo posee las competencias docentes requeridas para su ingreso laboral al sistema educativo oficial y durante su permanencia demuestra sus competencias para el cargo que ocupa*.(MEC 2008). El diseño ha sido construido participativamente con los actores del sistema tales como los gremios de educadores, docentes en servicio, responsables nacionales y locales de la educación; además se ha consultado con expertos internacionales. Su aplicación se halla en la fase inicial con la evaluación para el acceso al cargo.

Así, los resultados de la evaluación de los egresados de los Institutos de Formación Docente ha posibilitado pensar en la necesidad de enfocar el desarrollo de las políticas de calidad con la institucionalización de un mecanismo permanente de evaluación de las instituciones formadoras, el seguimiento del desempeño en el aula mediante la Certificación y el uso de toda esa información para el desarrollo del sistema de formación continua de los educadores.

Técnicamente este ha sido un proceso de mucho aprendizaje y participación en la construcción de los diseños, la validación, etc., sin embargo se requiere de un gran trabajo político en dos direcciones: (i) con los gremios docentes para instalar como agenda de estado y (ii) con el Congreso para obtener el financiamiento necesario. Hasta el momento los recursos utilizados fueron en buena parte créditos internacionales del BID y del Banco Mundial.

¹⁴ El MEC ha firmado un acuerdo con la agencia para llevar adelante el proceso de evaluación. Se puede consultar en www.aneaes.gov.py

¹⁵ Sobre los mecanismos de Licenciamiento y Acreditación y Evaluación del desempeño para la certificación se han publicado documentos que se citan en la bibliografía y también se puede consultar en www.mec.gob.py

2.4.2. Evaluación en la Educación inicial y preescolar

Desde el año 1995 se trabajó en la construcción de una política de primera infancia impulsado por la sociedad civil y el Estado con el apoyo de la cooperación internacional en el contexto de la implementación de la reforma de la educación escolar básica.

El Plan nacional tiene un enfoque de política intersectorial para el desarrollo integral de niños y niñas menores de 6 años con importante énfasis en la lucha contra la pobreza así como un papel muy relevante del sector educativo para la expansión de las oportunidades educativas dado que antes de la Reforma el acceso al pre escolar se daba solo en zonas urbanas y en instituciones de gestión privada.

La evaluación de la calidad de la educación preescolar es una de las líneas del Programa de Mejoramiento de la Calidad de la Educación Inicial y Preescolar desarrollado en el marco del Plan Nacional de Educación Inicial.(MEC 2002).

Por primera vez se evaluó la calidad del sistema preescolar, tomando en cuenta los aspectos de estructura, proceso y orientación y su influencia en el desarrollo infantil¹⁶.

Diseño metodológico¹⁷

El diseño abarcó una medición del desarrollo infantil al inicio y al final del preescolar. La muestra estuvo conformada por 100 salas de preescolar, 100 docentes y directores, 437 niños y niñas y sus respectivas familias (352 del grupo escolarizado y 85 del grupo control).(MEC 2008) documento interno que se halla en fase de impresión.

El diseño utilizado corresponde a un estudio cuasi experimental, como una medición al inicio de al año, una medición al final y con un grupo experimental y un grupo de control. También incluyó observaciones de la sala de preescolar y visitas a las familias de los niños y niñas, tanto del grupo experimental como de control.

Instrumentos aplicados

Se aplicaron instrumentos internacionales que permite estimar los efectos de la calidad de los entornos educativos (escuela-familia) sobre el desarrollo de los niños y las niñas de 5 años. Los instrumentos fueron sometidos a estudios de coherencia y pertinencia cultural. Se evaluaron tres constructos (i) **desarrollo infantil**, las variables estudiadas han sido destreza motora, lenguaje, pensamiento interacción con pares y adultos , (ii) **la calidad del ambiente escolar**, se consideró el clima de aula, la formación del educador, la gestión institucional y (iii) el **entorno familiar** abarcó variables tales como la calidad de la educación en el hogar y las características del ambiente.

Si bien el estudio no ha sido aún publicado hasta la fecha, debido al cambio de administración en el MEC, los resultados preliminares indican que la calidad en las salas no alcanzan los estándares

¹⁶ La recolección de datos se llevó a cabo durante el año 2005 y parte del 2006. El informe se halla en impresión. Las informaciones que se citan corresponden a documentos internos del MEC.

¹⁷ El diseño de la evaluación ha sido presentado y debatido en el Simposio Latinoamericano de Evaluación de la Calidad del Preescolar realizado en Asunción y sus resultados publicado en (MEC 2006) Modelos conceptuales y metodológicos en la Evaluación de la calidad de la educación pre escolar.

esperados. El estudio identificó que la intencionalidad pedagógica es débil lo cual estaría indicando la falta de conocimiento del desarrollo infantil.¹⁸

En el mismo estudio el análisis de factores asociados muestra que es el contexto familiar el que más incidencia tiene en el desarrollo infantil, tanto en aspectos cognoscitivos, lenguaje y socio afectivos.

Estos primeros pasos en la producción científica para el desarrollo conceptual y metodológico de la evaluación de la calidad del nivel inicial constituyen las bases para orientar las políticas públicas a favor del desarrollo infantil. El acceso de los niños y niñas al pre escolar es un primer avance, hacerlo con calidad es un reto y esto no será posible sin el soporte de los conocimientos.

2.4.3. Evaluación del rendimiento académico y del pensamiento crítico en la educación media

En la educación media podría decirse que se ha avanzado hacia una segunda generación de la evaluación. En primer lugar por el enfoque de los instrumentos orientados hacia la evaluación de competencias, el uso de bloques en la elaboración de pruebas y la incorporación del TRI en el análisis de los resultados en las áreas de Lengua y Literatura Castellana, Matemática y Ciencias Básicas. En segundo lugar porque además de la muestra nacional se ha podido dar seguimiento a un conjunto de instituciones beneficiarias de un proyecto focalizado con el objeto de monitorear las políticas de equidad y en tercer lugar el desarrollo y aplicación del test de pensamiento crítico. Con esto el SNEPE ya no sólo mide el aprendizaje del currículo, sino que avanza hacia la estimación del efecto del currículo en el desarrollo del pensamiento crítico. Los resultados han sido publicados en dos informes, uno sobre el rendimiento académico con la utilización del TRI y el otro sobre el pensamiento crítico que figuran en la bibliografía (MEC 2007). Junto con estos cambios también cabe mencionar que se evalúa a estudiantes del segundo curso de la educación media y ya no al tercero como se señalaba anteriormente. El cambio de curso para la evaluación es una medida de política considerando que los resultados podrían ser mejor utilizados con fines de mejoramiento durante la permanencia de los estudiantes en el colegio y no solo como información de salida del último año de este nivel educativo.

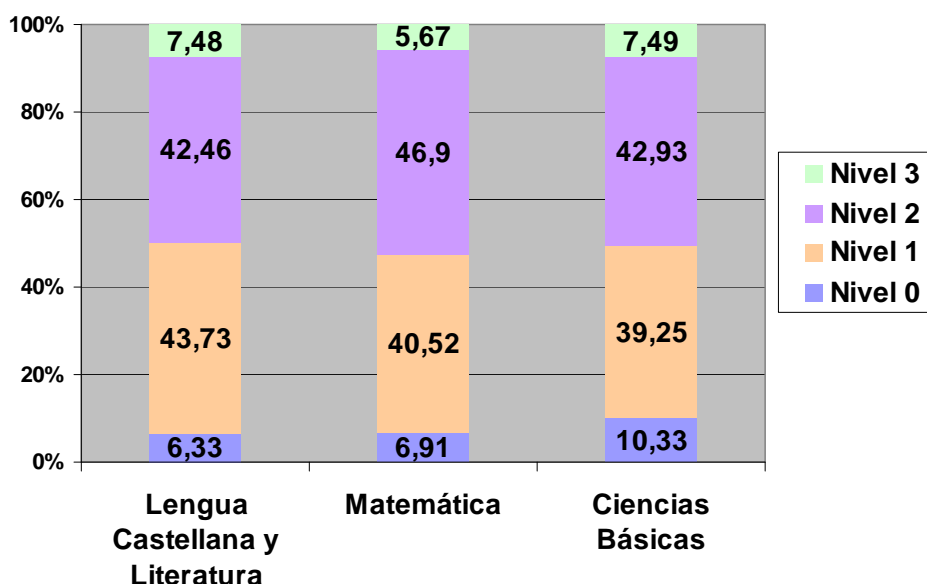
En el año 2006 es realizada la primera medición con el nuevo enfoque modelo de evaluación por competencias así como también fue aplicado el primer Test de Pensamiento Crítico. A modo de ejemplo se exponen algunos resultados con el nuevo enfoque.

Según el informe nacional (Snepe 2007) las pruebas resultaron difíciles para los jóvenes conforme al comportamiento del mapa de habilidades y dificultades. En Lengua y Literatura el mayor porcentaje de estudiantes se concentran en los niveles 1 (lectura literal) y 2 (lectura inferencial), en el nivel 3 (lectura inter textual) sólo el 7,48%.

En matemática ocurre algo similar dado que en el 5,67 % de los jóvenes se ubican en el nivel 3 (Inferencia y modelización de estrategias para resolver problemas complejos); alcanzan el nivel 2 el 46,90% de los estudiantes (conexión e integración de datos para resolver problemas de contextos familiares); en el nivel 1 el 40,52 % (reproducción y realización de operaciones sencillas).

¹⁸ (MEC 2008) Informe de resultados de la evaluación de la efectividad del pre escolar. Documento de circulación restringida.

GRÁFICO 6. LOGROS NACIONALES; PORCENTAJE DE ESTUDIANTES POR NIVELES DE DESEMPEÑO



Fuente: MEC, DGDE, SNEPE 2006/2007

La evaluación del pensamiento crítico en la educación media

La expectativa del desarrollo del pensamiento crítico constituye un propósito importante en los diferentes niveles del sistema educativo nacional. La Reforma Educativa de la Educación Escolar Básica lo menciona entre las competencias a lograr, en la educación media es una de las metas terminales del nivel y en el ámbito de la Educación Superior es un componente de los perfiles de salida de las carreras humanísticas o sociales.

El sistema educativo de la dictadura dejó muy asociada a la educación formal con el vaciamiento de su sentido crítico, por ello la demanda ciudadana ha sido muy clara en la necesidad de la formación del pensamiento crítico y creativo en el proceso de Reforma Educativa; se menciona esto como un contexto histórico que ha impactado los propósitos de la educación.

Ante un mundo signado por las desigualdades y en plena era del conocimiento se reconoce la relevancia que tiene el pensamiento crítico para el desenvolvimiento de las personas en diferentes ámbitos, muy especialmente en la producción y verificación del conocimiento científico, en la vida ciudadana tanto en su dimensión social como para la construcción de un proyecto de vida personal y en el de los procesos educativos formales para seguir aprendiendo. Sin embargo esta comprensión no siempre se corresponde con el desarrollo de estudios masivos en la educación pública con fines pedagógicos.

La reforma le ha dado mucho énfasis a la formación de la ciudadanía por lo cual el sistema de evaluación debía dar seguimiento a ese propósito con la incorporación de la medición del pensamiento crítico social. Este desafío ha sido asumido por el Snepe con el apoyo de expertos externos ya que hubo que elaborar desde el marco conceptual, preparar el test y validarlo.

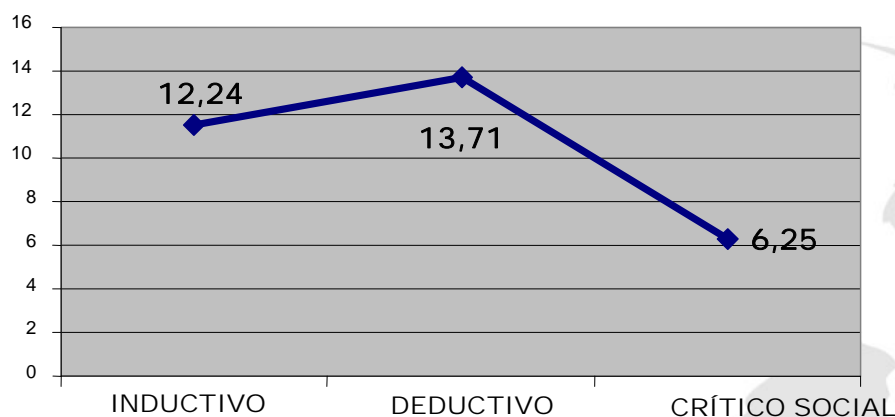
Se lo definió como *"la capacidad de develar, con su justificación, inconsistencias en un discurso formal, inferencias ilegítimas a partir de datos de la realidad natural o interpretaciones insuficientes, ingenuas o sesgadas de la realidad social"* (MEC2007)

Operacionalmente se define como la capacidad de un sujeto (alumno/a) de responder adecuadamente a un conjunto de tareas (ítems) que incluyen silogismos, identificación de supuestos, inferencia a partir de datos, distinción entre hechos y opiniones e interpretación de una "escena" (estímulos de tipo proyectivo) sobre una situación de carencia o conflicto. Se consideró que un estímulo proyectivo que representa una situación humana, individual o colectiva es más adecuado para suscitar "percepciones" que pueden interpretarse con un sentido valorativo. (MEC 2007)

El informe nacional (Snepe 2007) muestra que la media nacional en proceso inductivo alcanzó 12,4; en proceso deductivo 13,71 y en proceso crítico social 6,25. La escala es de 20 puntos para cada dimensión, a mayor puntaje mejor resultado. Según los cortes interpretativos de los componentes los procesos inductivo y deductivo muestran resultados comparables situándose en un nivel medio alto, sin embargo el proceso crítico social se ubica en un nivel medio bajo de criticidad. Aun no se ha establecido un estándar de desempeño para la educación media. Será necesario reunir mayor base empírica y discusión de políticas.

El siguiente gráfico es un ejemplo del resultado nacional. Se ha evaluado a una muestra de estudiantes del segundo año de la educación media. El trabajo ha sido publicado en un extenso informe que contiene el marco teórico sobre los estudios del pensamiento en la literatura internacional, el diseño metodológico y los resultados¹⁹.

GRÁFICO 7. APLICACIÓN 2006. MEDIAS GENERALES POR COMPONENTE
PUNTAJE MÁXIMO POR COMPONENTE: 20 PUNTOS



2.5. Evaluaciones y estudios internacionales

¹⁹ MEC (2007). Informe de Resultados Test de Pensamiento Crítico. Segundo Curso. Sistema Nacional de Evaluación del Proceso Educativo. Programa Reforma Joven .MEC-BIRF. Asunción. Paraguay.

El SNEPE se constituye en un punto focal para la participación del país en estudios y evaluaciones internacionales. Hasta ahora el MEC se halla en vinculación con tres proyectos que se describen a continuación.

2.5.1. Laboratorio Latinoamericano de Evaluación de la Calidad de la Educación. UNESCO/OREALC

La primera experiencia de evaluación internacional en la que participó el Paraguay es el Laboratorio Latinoamericano de Evaluación. Desde el año 1996 el Snepe forma parte del Laboratorio. El primer estudio se realizó en 1997 y el segundo en el 2006²⁰. Además de ser la primera experiencia de evaluación comparada el Laboratorio ha sido un espacio de formación y actualización muy importante para el desarrollo profesional de los técnicos del Ministerio de Educación. Ha permitido la interacción con expertos de todo el mundo y ha respondido a las necesidades de aprendizaje de los países. Además de evaluar permitió la formación de una masa crítica principalmente en el diseño de pruebas y análisis de resultados. A partir del SERCE el Laboratorio brindará apoyo pedagógico a aquellos países con resultados inferiores a la media regional.

2.5.2. International Civic and Citizenship Education Study (ICCS)²¹

El Estudio Internacional de Educación Cívica y Ciudadana es un proyecto del cual participan Chile, Colombia, Guatemala, México, Paraguay y República Dominicana ; está respaldado por Centro Regional para el Fomento del Libro en América Latina y el Caribe (CERLALC), el Sistema Regional de Evaluación y Desarrollo de Competencias Ciudadanas (SREDECC) Asociación Internacional para la Evaluación de Logros Educativos – IEA

Tiene como propósitos contribuir a la construcción de una ciudadanía democrática en los países de América Latina mediante el fortalecimiento de la educación para la ciudadanía y la democracia e investigar la preparación de los jóvenes para asumir su rol de ciudadano.

Aborda contenidos tales como sociedad civil y sistemas, principios cívicos, participación cívica e identidad cívica. En la dimensión afectivo conductual investiga sobre valores, actitudes, intenciones y conductas; en la dimensión cognitiva sobre conocimiento, razonamiento y análisis.

Los instrumentos utilizados son (i) prueba Internacional Cognitiva, (ii) cuestionario Internacional sociodemográfico de estudiantes (iii), cuestionario internacional sobre percepción de estudiantes, (iv) prueba cognitiva regional, (v) cuestionario a estudiantes regional, (vi) cuestionario a docentes, (vii) cuestionario a escuelas y (viii) encuesta sobre contexto nacional.

En el año 2008 los instrumentos han sido validados y en el 2009 se tiene prevista la evaluación a la muestra internacional seleccionada.

Esta participación también apoya el proceso de aprendizaje de los técnicos del MEC.

²⁰ <http://www.llece.unesco.cl>

²¹ Mayor información se encuentra en la web de la IEA www.iea.nl

2.5.3. Literacy Assessment and Monitoring Program LAMP. Programa de Evaluación y Monitoreo de la Alfabetización de adultos. Coordinado por el Instituto de Estadísticas de la UNESCO (UIS) Montreal, Canadá

El Ministerio de Educación con el apoyo del Ministerio de Educación de España, la Cooperación Española y de la Organización de Estados Iberoamericano para la Educación, la Ciencia y la Cultura, OEI vienen ejecutando un plan de Alfabetización y Educación Permanente de Jóvenes y Adultos. Se iniciaron los primeros pasos para la instalación de un sistema de evaluación de la calidad de la educación de adultos *con base en los niveles de desarrollo*. Se gestiona la formalización del acuerdo mediante la firma de un memorando de entendimiento entre Paraguay y UNESCO. Según el cronograma se tiene pensado la prueba piloto del sistema en el primer semestre del año 2009.

Dos de los principales requisitos que establecen el valor de cualquier acción son, en primer lugar, el que ella sea lo que dice ser, en otras palabras si es o no **efectiva** y si acaso como tal cumple con su propósito, en otros términos si es o no **eficaz**. De paso, es importante señalar que de hecho, la **efectividad** condiciona la **eficacia**.

Sin embargo, hay pocos términos en el campo de la educación cuyo sentido sea confundido con mayor frecuencia que el de estos dos. Se habla, por ejemplo, de “escuelas efectivas” cuando se quiere hacer referencia a “escuelas eficaces” y todo ello debido a una errónea traducción desde el idioma Inglés, del cual nos llegan la mayoría de las expresiones técnicas en educación. En efecto, tendemos a confundir los términos, ya que por una parte, la traducción correcta del término inglés “effectiveness” al castellano es **eficacia** -lo distintivo de aquello que cumple con su propósito- y por la otra, el término “efectivo” en castellano describe a aquello que es real.

3. EL PAPEL DEL SNEPE EN LA REFORMA: APRENDIZAJES

Hasta aquí se ha intentado realizar una síntesis del contexto nacional y el desarrollo del sistema de evaluación como una de las políticas de la Reforma Educativa desde su creación hasta el año 2008. En esta sección se reflexiona sobre los aprendizajes de la experiencia y se plantean los desafíos para el futuro en un contexto de cambio político en el país.

Desde el segundo semestre del año 2008 el Paraguay vive un nuevo proceso político con el cambio histórico resultante de las elecciones del 20 de abril con el triunfo de la Alianza de partidos de oposición y movimientos sociales luego de 60 años de gobierno del partido colorado. La Reforma Educativa por primera vez se enfrenta con las incertidumbres de su futuro teniendo en cuenta que durante sus años de implementación ha sido considerada como la única agenda de Estado de la transición con el desarrollo del Plan Estratégico Paraguay 2020 desde la década de los 90, la estabilidad y el compromiso de un cuadro técnico del MEC por más de una década. Entre los principales aprendizajes de este proceso, cabe mencionar los siguientes:

- La voluntad política es fundamental para un sistema de evaluación. Instalar y mantener un sistema de evaluación requiere de un decidido compromiso de las máximas autoridades. La experiencia ha sido positiva y la clave ha estado en la comunicación permanente del equipo directivo del sistema con sus autoridades lo cual ha posibilitado el involucramiento de los ministros y ministras de educación principalmente en la difusión de los resultados que constituye el momento de mayor visibilidad y participación de los niveles superiores de la organización.

- La conformación de un equipo comprometido y técnicamente riguroso. En un país sin antecedentes de evaluación de sistema hubo que identificar a los técnicos y formarlos. Se trabajó con un equipo mixto integrado por técnicos del MEC y consultores nacionales permanentes para la elaboración de instrumentos, el procesamiento, la elaboración de los informes y la difusión. Las consultorías internacionales se contrataban para la supervisión del proceso y principalmente para el análisis de la información. La norma de todo apoyo externo nacional o internacional siempre fue la formación y consolidación del equipo para la sostenibilidad. Esto ha actuado como un factor de motivación importante para el equipo. Otro aspecto relevante para el aprendizaje ha sido el contacto con otros sistemas de la región en agendas bilaterales y la participación en el LLECE. Esa ha sido una escuela para los profesionales del MEC.
- La confianza y credibilidad de un sistema de evaluación es capital. En un país con alta percepción de corrupción y de desconfianza hacia el servicio público demandaba del SNEPE maximizar los cuidados en la confidencialidad de los instrumentos durante su elaboración e impresión y en la selección de los examinadores. Ningún procedimiento podía quedar librado a la discrecionalidad. Se implementaron protocolos para la selección y entrenamiento de los examinadores y un proceso de supervisión permanentes. Se desplegaba un gran esfuerzo de control con el personal de las imprentas. Pero lo más efectivo ha sido construir alianzas y compromisos con todos los involucrados para preservar la confidencialidad. No se trataba solamente de operar con normas extrínsecas sino de ganar aliados para una tarea compartida y de un interés superior para el país. Se puede decir que la experiencia resultó exitosa porque no se conoce de casos de filtración de instrumentos. Tampoco se ha puesto en duda la credibilidad en los resultados.
- La difusión oportuna de la información y la transparencia. No basta que las informaciones sean válidas y confiables si no son oportunas. Cumplir con los cronogramas de divulgación ha ayudado a la credibilidad y cualquier retraso por motivos técnicos, que siempre existen, ha sido interpretado como falta de transparencia principalmente por los medios de comunicación. A lo largo de los años se ha cumplido con los plazos. Sin embargo en el año 2007 ha habido problemas con la gestión para la adjudicación de la imprenta lo cual ha generado casi una interpelación a las autoridades por parte de los medios. Si bien es una situación incómoda cuando se está en la gestión pensamos que es un cambio positivo en la cultura de la exigencia de rendición de cuentas públicas. Una lección importante es que si se ya se sabe que se tendrán retrasos hay que adelantarse y anunciar para conservar ese capital tan sensible que es la confianza.
- La densidad y multiplicación de la agenda de aplicación de pruebas. Es muy positivo el avance dado en cuanto a la expansión de los niveles evaluados, la incorporación de nuevas áreas y sobre todo las innovaciones como la evaluación del pensamiento crítico y el trabajo de la educación inicial. Sin embargo esto tiene su contra cara en cuanto una inmensa cantidad de datos que ya no alcanzan a ser profundizados. Mucho operativo de evaluación e insuficiente análisis respecto al aprovechamiento de datos.
- Los requerimientos técnicos de alta especialización. Es sabido que los sistemas de evaluación requieren del concurso de profesionales de alto nivel por los saberes técnicos necesarios. Como ya se señaló el SNEPE fue hallando soluciones al tema contratando expertos externos para el apoyo al equipo de planta. Sin embargo un equipo sólido requiere de la permanencia de esos

perfiles profesionales y existe dificultad financiera para retener profesionales de alta calificación con rubros del estado por no resultar atractivos y por estar por debajo de las expectativas. Las implicancias son la dependencia de recursos externos, la falta de posibilidad de elevar la calidad y complejidad científica de los trabajos.

- La difusión y el uso de los resultados. La experiencia indica que la producción de información de bien público, la difusión y promoción del uso más diverso y amplio posible excede las competencias de un equipo de evaluación. Como se analiza en la literatura internacional esta es una fase o área débil de los sistemas. Los informes se reducen a reproducir cuadros y gráficos con comentarios con un lenguaje más numérico que interpretativo lo cual reduce el impacto que pudiera tener la información. Esto ha provocado discusiones internas con el equipo de comunicación social de uno de los programas de la reforma. A partir de ello algunos resultados fueron publicados en una revista pedagógica con llegada a todas las escuelas del país con un lenguaje más pedagógico, contextualizado y de edición trimestral. El problema está abierto y requiere nuevas decisiones. Algunos investigadores²² lo consultan aunque esa práctica no es frecuente. La cooperación internacional la consulta para los estudios sectoriales y para el diseño de proyectos de mejoramiento. La difusión interna a otras dependencias del MEC ha mejorado principalmente en formación docente, currículo, elaboración de textos y materiales. A nivel de aula y opinión pública falta un largo camino por recorrer.
- Los factores asociados y los resultados del aprendizaje. Si bien se recogen muchas informaciones el análisis es insuficiente y sobre todo falta una mejor identificación de las variables explicativas de los resultados tanto aquellas que se vinculan a aspectos internos de la escuela y el peso de las variables de contexto. Hasta ahora el trabajo es muy descriptivo y no da sustento para generar nuevas políticas públicas, lo que reduce la posibilidad de elaborar sugerencias pedagógicas y metodológicas con base a un sistema de información basado en evidencia empírica.

4. EL CAMBIO POLÍTICO EN EL PAÍS Y LA NUEVA GESTIÓN DEL MEC

El cambio político supuso también el cambio en la administración y gestión del Ministerio de Educación. En el periodo de transición, mayo a agosto del 2008, los equipos salientes y entrantes han mantenido reuniones de trabajo con el objetivo de hacer un cambio lo mas institucional posible. En ese lapso uno de los temas abordados fue el SNEPE, su situación y futuro. Hubo coincidencia sobre la necesidad de un mayor fortalecimiento y desarrollo en base a las lecciones aprendidas y a los desafíos que debía enfrentar.

Un primer aspecto ha sido su ubicación institucional como un componente central del sistema de información. Al respecto la nueva administración ha decidido conservar la Dirección de Evaluación pero con cambio de su dependencia. Pasó a formar parte de la Dirección General de Planificación Educativa. Actualmente se encuentran en plena etapa de transición y no se dispone de datos que pudieran permitir un análisis más profundo. Como hipótesis el cambio puede ser positivo si se logra consolidar los dos

²² Al respecto Ernesto Schiefelbein ha realizado un estudio sobre Educación y Empleo en el Paraguay utilizando como una de las fuentes el SNEPE y está publicado en (CADEP 2007). Asunción Paraguay.

sistemas, el de información de las estadísticas y el de evaluación siempre y cuando se distingan los requerimientos técnicos y las diferencias entre ambos. La idea es buena, el paso es oportuno y el futuro dirá sobre sus resultados.

4.1. El futuro y sus desafíos

Una sociedad más crítica y con altas expectativas sobre el valor de la educación, exigirá un mayor ensanchamiento de la discusión pública, democrática y bien informada sobre la calidad de la educación, en un país que cree en la educación como su único camino para la movilidad social en una realidad de pobreza y desigualdad

Ese escenario tiene varias implicancias, aquí identificamos algunas:

- **Financiamiento:** hasta el momento el sistema de evaluación se financia principalmente mediante los créditos externos. Una de las razones es que el presupuesto de la educación en más del 90% se utiliza para gastos corrientes y no se ha logrado el cambio de las fuentes externas a los fondos genuinos del tesoro. Los salarios de los funcionarios se financian por fuente nacional, sin embargo los insumos, operativos y consultores se financian a través de préstamos. En años anteriores se ha intentado aplicar una estrategia de financiamiento compartido y gradual sin embargo a la hora de la disponibilidad presupuestaria tales recursos no se disponían.
- **Formación continua del equipo e innovación:** el avance científico permanente de la evaluación y la complejidad de los análisis requeridos demandan una mayor solvencia profesional para la elaboración y adaptación de instrumentos, la capacidad analítica e interpretativa de los resultados y el aprovechamiento de toda la información que se recoge.
- **Difusión y uso de la información:** demanda nuevas competencias para la elaboración de informes y difusión de los resultados con fines de educativos y pedagógicos. Resulta repetitivo y poco efectivo el tratamiento cuantitativo de los datos sin profundidad en las interpretaciones y con muy escasas orientaciones pedagógicas para transformar las prácticas de aula que es en verdad lo que importa. Las tipologías de informes cumplieron su ciclo, ya mostraron lo que pueden dar y sobre todo sus limitaciones. Tuvimos que pasar por ello, más lo que no se puede es provocar saturación, rechazo y desperdiciar oportunidades para el mejoramiento. La pregunta institucional es si este proceso debe quedar a cargo del sistema de evaluación o si acaso deba ser encarado por la instancia de comunicación social. Más allá del marco institucional lo que debe comprenderse es la necesidad de convertir la información evaluativa como orientación para el mejoramiento pedagógico. Como sea requiere nuevas competencias y creación de interfases de equipos transdisciplinarios.
- **Avanzar en la construcción de estándares.** La pregunta de la ciudadanía silenciosa y a veces explícita es si "estamos bien o mal" cada vez que se publican resultados. Es cierto que el uso del TRI en parte ayuda sin embargo ya es tiempo de proponerse tal meta, las condiciones ya están dadas. Las reflexiones hechas hasta ahora dejan ver que no será conveniente disponer de un solo parámetro porque la evidencia muestra que existe mucha distancia entre los resultados. Por ello será más estimulante fijar parámetros que permitan conseguir calidad gradualmente.
- **Consolidar un equipo para la participación y seguimiento de los estudios internacionales.** El actual equipo se encuentra excedido de responsabilidades por la amplitud y diversidad temática de la agenda nacional de evaluación, lo cual limita la plena participación y mejor

aprovechamiento de las experiencias. Además se deberá considerar la incorporación del Paraguay a Pisa por los beneficios formativos. Esta posibilidad ha sido explorado y por razones presupuestarias no se ha podido efectivizar.

- Meta evaluación del sistema de evaluación: será oportuno constituir un equipo mixto integrado por profesionales nacionales y expertos internacionales y encarar una revisión de los logros, construir una priorización de la agenda futura con estimación de costos y fuentes de financiamiento. Por la salud del sistema educativo puede ser muy potente contar con un apoyo de esta naturaleza con el objeto de posicionar al Snepe como un mecanismo confiable, válido y con pertinencia para la toma de decisiones.
- Uso de bases de datos y resultados de evaluación por académicos: el uso ha sido escaso disponiendo de tanta riqueza de información. En parte se debe a poca promoción del propio sistema y por otra al desconocimiento y escasez de instancias de investigación en el país. Se requieren muchos estudios cualitativos que profundicen y den seguimientos a los hallazgos y a las hipótesis.
- Construir una estrategia de trabajo permanente con los medios de comunicación: los medios de comunicación logran fuerte impacto en la opinión pública y la experiencia del tratamiento de los informes de evaluación requiere de una política intencionada y de la construcción de una agenda de trabajo permanente. Ojala se logre consensuar una agenda básica de la información como bien público sin pretender que desaparezca el uso interesado de los resultados según la agenda de los medios. Pero al menos mitigar para la democratización de una información tan sensible como la calidad de la educación en favor del empoderamiento y la participación de la sociedad civil.

En suma, el MEC logró diseñar y desarrollar un sistema de evaluación que ha acompañado y detectado las fortalezas y debilidades de la política educativa. Es un sistema bien posicionado ante la ciudadanía, los educadores, los gremios de educadores, los medios de comunicación. La voluntad política de las autoridades, el rigor técnico, el compromiso, la estabilidad del equipo, la participación en espacios internacionales y el apoyo de la cooperación internacional han favorecido la confianza y credibilidad del que goza. Su efectividad y pertinencia dependen de renovadas voluntades, más recursos profesionales, tecnológicos y financieros y un perseverante proceso de aprendizaje de nuevas capacidades.

REFERENCIAS BIBLIOGRÁFICAS

- Banco Mundial (2008). *Educación Media en el Paraguay. Logros, opciones y desafíos*. Informe N° 42665-PY. Washington: Banco Mundial.
- Centro de Análisis y Difusión de la Economía Paraguaya, CADEP (2007). *Borda, Dionisio. Editor. Economía y Empleo en el Paraguay*. Asunción: CADEP.
- Consejo Asesor de la Reforma Educativa (1992). *Informe de Avance Fundación en Alianza*. Asunción: CARE.
- Dirección General de Estadísticas, Encuestas y Censo (2003). *Censo Población y Vivienda 2002*. Asunción: DGEEC.
- Laboratorio Latinoamericano de Evaluación de la Calidad de la Educación. LLECE (2008). Segundo Estudio

- Regional Comparativo y Explicativo, SERCE.UNESCO. OREALC. Santiago de Chile: UNESCO.
- Marchesi, A y Martín E.(2007). *Evaluación del Sistema Educativo de Paraguay*. Asunción: Organización de Estados Iberoamericanos para la Educación, la Ciencia y la Cultura.
- MEC (1988). *Ley General de Educación*. Asunción: MEC.
- MEC (1995). *Propuesta para el Sistema Nacional de Evaluación del Proceso Educativo*. Asunción: MEC.
- MEC (1997). *Primer Informe de resultados. Sexto Grado. Sistema Nacional de Evaluación del Proceso Educativo*. Programa de Mejoramiento de la Calidad de la Educación Primaria.. Asunción: MEC-BID.
- MEC/CARE. (1998). *La Reforma Educativa Avanza*. Asunción: MEC/CARE.
- MEC- Universidad Católica Nuestra Señora de la Asunción. (1999). *Estudio sobre educación media y desarrollo en el Paraguay*. Programa MECES. Asunción: MEC.
- MEC (1999). Resultados de las Pruebas Nacionales. Sexto Curso. Sistema Nacional de Evaluación del Proceso Educativo. Programa de Mejoramiento de la Calidad de la Educación Secundaria. Asunción: MEC-BIRF.
- MEC (2001). *Un horizonte nuevo para los Institutos de Formación docente*. Asunción: MEC.
- MEC (2001). *Informe de las Pruebas Nacionales. Tercer Curso y Sexto Curso. Sistema Nacional de Evaluación del Proceso Educativo*. Programa de Mejoramiento de la Calidad de la Educación Secundaria. Asunción: MEC-BIRF.
- MEC (2002). *Cuanto aprenden nuestros niños y niñas. Informe de resultados. Tercer Grado y Sexto Grado. Sistema Nacional de Evaluación del Proceso Educativo*. Programa Escuela Viva Hekokatúva. Asunción: MEC-BID.
- MEC (2002). *Plan Nacional de Educación Inicial y Preescolar*. Asunción: MEC.
- MEC (2002), (2004) y (2008). *Informe de Resultados. Cómo redactan los futuros maestros*. Asunción: MEC.
- MEC (2003). *Manual de procedimientos para la Autoevaluación de los Institutos de Formación Docente*. Asunción: MEC.
- MEC (2004). *Sistematización de experiencias pedagógicas de instituciones de formación docente*. Documento interno Dirección de Formación Docente.
- MEC (2006). *Modelos Conceptuales y Metodológicos en la Evaluación de la Calidad de la Educación Preescolar*. Programa Escuela Viva Hekokatúva. MEC-BID. Asunción: MEC.
- MEC (2007). *Informe de Resultados de la educación Media. Segundo Curso. Sistema Nacional de Evaluación del Proceso Educativo*. Programa Reforma Joven .MEC-BIRF. Asunción. Paraguay
- MEC (2007). *Informe de Resultados Test de Pensamiento Crítico. Segundo Curso. Sistema Nacional de Evaluación del Proceso Educativo*. Programa Reforma Joven .MEC-BIRF. Asunción: MEC.
- MEC (2007). *Mecanismo de Licenciamiento de Instituciones Formadoras de Docentes*. Asunción: MEC.
- MEC (2008). *Resultados preliminares estudios del preescolar*. Documento de circulación restringida.
- MEC (2008). *Evaluación del Desempeño para la Certificación Profesional de los Educadores del Paraguay*.

Proyecto de Reforma de la Educación con Énfasis en la educación Media. Asunción: MEC.

Reimers, F y Mc Ginn, N. (1997). *Diálogo Informado*. Centro de Estudios Educativos. México DF

Rivarola, D. (2000). *La Reforma Educativa en el Paraguay*. Santiago de Chile: CEPAL.

Paginas WEB:

<http://www.arandurape.edu.py>

<http://www.aneaes.gov.py>

<http://www.llece.unesco.cl>

<http://www.mec.gob.py>

<http://www.pisa.oecd.org>

<http://www.iea.nl>



LA CALIDAD DEL SISTEMA EDUCATIVO MEXICANO DESDE LOS RESULTADOS DE EVALUACIONES NACIONALES. EL APRENDIZAJE EN MATEMÁTICAS

Guadalupe Ruiz

Revista Iberoamericana de Evaluación Educativa 2009 - Volumen 2, Número 1

<http://www.rinace.net/riee/numeros/vol2-num1/art4.pdf>

Fecha de recepción: 15 de marzo de 2009

Fecha de comunicación de dictamen: 20 de abril de 2009

Fecha de aceptación: 20 de abril de 2009

En México la evaluación externa en gran escala se remonta a los años setenta del siglo pasado con el desarrollo de proyectos que fueron de notable envergadura en su momento. Sin embargo, no es sino hasta la década de los noventa que la evaluación adquiere un estatus distinto, caracterizado por una mayor presencia, relevancia y reconocimiento, lo cual se asoció, en buena medida, a la descentralización educativa que se impulsó en el país a partir de 1992 y a otras medidas de política educativa (programas compensatorios, estímulos salariales a los docentes, programas de actualización y acreditación) que incluyeron un importante componente evaluativo. Pese a ello, una constante a lo largo de este amplio periodo fue la nula difusión de los resultados de la evaluación y su escaso aporte a la fundamentación de la toma de decisiones y la política educativa (Ruiz, 2003).

En el marco de una nueva administración federal en el país, en 2002 fue creado el Instituto Nacional para la Evaluación de la Educación (INEE), con la misión de evaluar la calidad del sistema educativo mexicano y hacer difusión oportuna de los resultados para contribuir a la mejora pedagógica, la toma de decisiones y la rendición de cuentas.

Este artículo se refiere a las evaluaciones sobre la calidad del sistema de educación básica en México realizadas por el INEE entre 2005 y 2007, años en los que se cubrieron los tres niveles que conforman la educación básica en el país (preescolar, primaria y secundaria). Son éstas, las evaluaciones más recientes realizadas en el país, que han tenido la pretensión expresa de evaluar la calidad del sistema en su conjunto y sus grandes subconjuntos y cuyos resultados han sido objeto de amplia difusión. Aunque estos resultados permiten, con ciertas reservas, una aproximación a las tendencias de la calidad de la educación en el país, no es éste el enfoque con que se abordan en el presente trabajo (ver un análisis de esta naturaleza en INEE, 2008); aquí interesa particularmente, atendiendo a cada una de las evaluaciones en lo individual pero también en conjunto, mostrar la desigual distribución de los resultados de aprendizaje en los distintos contextos socioescolares en que tiene lugar la educación básica.

Dos áreas académicas típicamente abordadas en las evaluaciones nacionales e internacionales del logro escolar, son las que en términos curriculares corresponden a las asignaturas de Español y Matemáticas, o bien, a los ámbitos de las habilidades o competencias comunicativas, de expresión escrita y comprensión lectora, y de desarrollo del pensamiento matemático. Este trabajo centra su atención en los resultados en el área de Matemáticas básicamente por dos razones: la mayor importancia relativa de la escuela, en comparación con la familia, en la adquisición de conocimientos y el desarrollo de habilidades en este campo; y, la necesidad de brindar mayor atención a un área que no ha sido objeto con la misma intensidad (como la lectura, la comprensión lectora y las competencias comunicativas en general) de políticas educativas y de intervención tendentes a mejorar los niveles con que se logran los propósitos formativos de la educación básica.

El artículo está integrado por cuatro apartados: en el primero se hace una somera descripción de los servicios educativos que constituyen los grandes subconjuntos de contraste del sistema educativo nacional en cada uno de los niveles que integran la educación básica; el segundo presenta información sobre las muestras en las que se hizo cada evaluación; el tercero refiere brevemente las principales características de los instrumentos de evaluación utilizados, y el cuarto presenta una selección de resultados de las evaluaciones que han tenido lugar entre 2005 y 2007 en preescolar, primaria y secundaria en México. Se incluye un apartado de reflexiones finales.

1. LA OFERTA DE EDUCACIÓN BÁSICA EN MÉXICO. UNA BREVE DESCRIPCIÓN DE SU DIVERSIDAD

La educación básica en México es ofrecida predominantemente por el sistema público. Al inicio del ciclo escolar 2006-2007, 91% de la matrícula era atendida en escuelas públicas (INEE, 2007b). En cada uno de los niveles educativos que integran la educación básica existen variantes en el servicio conocidas como modalidades. En preescolar y primaria son esencialmente las mismas: general, educación indígena y educación comunitaria. En secundaria existe también la modalidad general, junto con la técnica y la telesecundaria que en conjunto atienden a 99% de la matrícula (INEE, 2007b). La tabla 1 ofrece una primera visión de esta oferta diversa.

TABLA 1. MODALIDADES POR NIVEL EDUCATIVO Y PORCENTAJE DE LA MATRÍCULA QUE ATIENDEN (CICLO 2006-2007)

Modalidad	Nivel educativo		
	Preescolar (%)	Primaria (%)	Secundaria (%)
General	89	93	50
Indígena	8	6	
Comunitaria	3	1	
Técnica			28
Telesecundaria			20
Otras			1

FUENTE: INEE (2007b)

En todos los niveles educativos la matrícula es atendida principalmente por la modalidad general. En preescolar y primaria, 89 y 93% de los alumnos respectivamente, acudía al inicio del ciclo escolar 2006-2007 a instituciones escolares clasificadas bajo esa modalidad (INEE, 2007b). La combinación de esta característica con el tipo de sostenimiento de la escuela –público o privado- y con su ubicación geográfica –en el medio urbano o rural- da lugar a una clasificación de escuelas que, para los efectos de la evaluación externa de los aprendizajes en el país configura algunos de los estratos de reporte de los resultados: urbano público, urbano privado y rural público.

La educación indígena de acuerdo a datos del mismo ciclo escolar, atendía a 8 y 6% de la matrícula escolar en preescolar y primaria respectivamente (INEE, 2007b). Este modelo educativo, bilingüe e intercultural, se ha desarrollado para responder en forma más pertinente que desde la modalidad general, a la diversidad cultural y lingüística de México. La educación indígena se concentra notablemente en poblaciones rurales (de menos de 2,500 habitantes): de acuerdo a datos del ciclo escolar 2004-2005, 90 y 93% de las escuelas de educación preescolar y primaria respectivamente, se encontraban asentadas en localidades de ese tamaño (INEE, 2005).

A inicios del ciclo escolar 2007-2008 las entidades con mayor matrícula en educación inicial y básica indígena eran cinco: Chiapas, Oaxaca, Guerrero, Puebla y Veracruz que en conjunto atendían al 72% de la matrícula de la modalidad (DGEI, 2007). Este sector de la oferta educativa da lugar a un cuarto estrato de reporte de resultados de la evaluación externa de aprendizajes: la educación indígena. Hasta el momento sólo se ha incluido en las evaluaciones realizadas en primaria (utilizando las mismas pruebas que en los demás estratos, es decir, en español), dado que en secundaria es prácticamente inexistente la modalidad, y en preescolar, aunque con una mayor presencia relativa, supondría la adaptación de los instrumentos de evaluación a diversas lenguas indígenas.

Finalmente, en el caso de la educación preescolar y primaria, la educación comunitaria constituye un modelo a través del cual se presta atención en escuelas unitarias a poblaciones aisladas y de reducido

tamaño. En estos servicios, un instructor sin formación profesional atiende a un número reducido de alumnos de todos los grados que abarca el nivel educativo, los que, en su mayoría pertenecen a los estratos más desfavorecidos de la población. En el ciclo 2004-2005 poco más del 90% de los cursos comunitarios de ambos niveles se encontraba asentado en localidades de entre 1 y 499 habitantes (INEE, 2005), esto es, aisladas o dispersas, carentes de servicios y por lo general con índices elevados de marginación. La educación comunitaria constituye el quinto estrato de interés en las evaluaciones nacionales de la calidad de la educación preescolar y primaria en México.

La educación secundaria como ya se ha dicho, se ofrece a través de tres modalidades principales: la secundaria general, la técnica y la telesecundaria. Las dos primeras son las más antiguas y predominantemente urbanas (87 y 75% de las secundarias generales y técnicas, respectivamente, se encontraban en localidades de ese tipo al inicio del ciclo escolar 2004-2005 (INEE, 2005). El plan de estudios de la modalidad general es el que siguen en su gran mayoría las escuelas privadas que ofrecen educación en este nivel. Las telesecundarias, por su parte, responden a un modelo a través del cual el sistema educativo mexicano ofrece educación secundaria a jóvenes que viven en comunidades rurales pequeñas y marginadas, en donde, por el reducido tamaño de la población escolar resulta incosteable establecer secundarias generales o técnicas. En estas secundarias un solo profesor es responsable de la enseñanza de todas las asignaturas en un mismo grado y se introducen recursos didácticos y tecnológicos especiales (INEE, 2007a).

Los cuatro estratos que han abordado las evaluaciones de resultados de aprendizaje en secundaria han sido pues los señalados: secundaria general, distinguiendo entre públicas y privadas, secundaria técnica y telesecundaria.

En todos los niveles de la educación básica, por lo ya dicho, la oferta diversa que proporciona el sistema educativo tiende a asociarse a una demanda también diferenciada. De esta suerte, y dado que, como bien ha demostrado la investigación sobre factores asociados al aprendizaje y eficacia escolar, los resultados de aprendizaje dependen en grado elevado de las condiciones de origen social de los educandos, todas las evaluaciones reportan consistentemente el impacto del entorno de los alumnos en su logro escolar y grandes brechas entre una y otra modalidad.

2. LAS MUESTRAS EVALUADAS

Entre 2005 y 2007 las evaluaciones del INEE cubrieron por primera vez los cuatro grados de la educación básica previstos en su Plan de Evaluación de Aprendizajes (3° de preescolar, 3° y 6° grados de primaria y 3° de secundaria). De acuerdo a lo señalado en el apartado 1 de este artículo, la misión de evaluar al sistema educativo mexicano en su conjunto ha obligado al INEE a definir muestras representativas cuyos resultados permitan hacer inferencias para el conjunto de la población (véase tabla 2).

Las dimensiones de las muestras de primaria y secundaria han permitido reportar resultados a varios niveles de agregación de la información: nacional, por estrato escolar, por entidad federativa y, en varios casos, por estrato y entidad simultáneamente. El caso de preescolar es distinto pues debido a la naturaleza especial de los instrumentos diseñados para la evaluación en ese nivel (ver apartado 3 de este documento), los niveles de reporte para los que fue construida la muestra abarcaron: el nivel nacional y el de los cuatro estratos escolares abordados en esta evaluación: urbano público, urbano privado, rural público y preescolar comunitario.

TABLA 2. CANTIDAD DE ESCUELAS Y ALUMNOS EVALUADOS EN PREESCOLAR, PRIMARIA Y SECUNDARIA (2005-2007)

Nivel y grado educativo	2005		2006		2007	
	Escuelas	Alumnos	Escuelas	Alumnos	Escuelas	Alumnos
3° preescolar					1,091	10,305
3° primaria			3,167	55,312		
6° primaria	2,770	47,858				
3° secundaria	2,397	52,251				

Fuente: Backhoff y otros (2006, 2007 y 2008).

En este artículo se presentan los resultados relativos sólo al nivel nacional y al de los estratos escolares. En los casos en que las evaluaciones fueron representativas de las entidades federativas, los hallazgos correspondientes pueden verse en los informes de dichas evaluaciones (Backhoff y otros, 2006, 2007 y 2008).

3. LOS EXÁMENES UTILIZADOS EN LAS EVALUACIONES DE LA CALIDAD DEL SISTEMA DE EDUCACIÓN BÁSICA

Las pruebas desarrolladas en el INEE para evaluar la calidad del sistema educativo nacional se conocen como Exámenes de la Calidad y el Logro Educativos, Excale. Entre sus características principales destacan su alineación al currículum nacional, su diseño matricial y su enfoque criterial. El hecho de que en México exista un único currículum para cada uno de los niveles que integran la educación básica (preescolar, primaria y secundaria) favoreció que en su momento se tomará la decisión de diseñar los Excale alineados a los planes y programas de estudio vigentes.

El aporte fundamental del diseño matricial es que da resultados confiables sobre una gama muy amplia de áreas, temas y contenidos curriculares, sin necesidad de someter a los alumnos a largas jornadas de resolución de pruebas, ya que una vez que se construye un conjunto de reactivos que cubra el dominio curricular que interesa evaluar, se divide en subconjuntos que se distribuyen entre los estudiantes, de tal manera que cada uno contesta sólo parte del examen completo. Este tipo de pruebas es especialmente apropiado para evaluar la calidad de un sistema educativo —que es lo que pretenden las evaluaciones a que hace referencia este artículo—, pero no para evaluar escuelas o alumnos en lo individual, quienes, como ya se ha dicho, responden una fracción de lo que en realidad es una prueba extensa. Entre otras razones, por ésta es que los resultados de estas evaluaciones no se reportan a nivel escuela, aula o alumno.

Los Excale, instrumentos criterios, fueron diseñados de forma que permitiesen reportar resultados en términos de los niveles de logro de los alumnos atendidos por el sistema educativo mexicano (véase tabla 3). Los niveles son definidos de acuerdo a una compleja y rigurosa metodología que involucra una gran variedad de especialistas —especialistas en currículum, en investigación educativa, profesores en servicio— en las distintas fases contempladas en el modelo para la determinación de niveles de logro (Jornet y Backhoff, 2006).

TABLA 3. DESCRIPCIÓN GENÉRICA DE LAS COMPETENCIAS ACADÉMICAS QUE LOGRAN LOS ALUMNOS EN CADA NIVEL DE LOGRO EDUCATIVO

Nivel	Descripción
Por debajo del básico	Indica carencias importantes en el dominio curricular de los conocimientos, habilidades y destrezas escolares que expresan una limitación para progresar satisfactoriamente en la materia.
Básico	Indica un dominio fundamental o elemental de conocimientos, habilidades y destrezas escolares necesarios para progresar satisfactoriamente en la materia.
Medio	Indica un dominio adecuado de conocimientos...
Avanzado	Indica un dominio óptimo que refleja el aprovechamiento máximo de lo previsto en el currículum.

FUENTE: Backhoff y otros (2006).

Los exámenes diseñados para primaria y secundaria son autoaplicados –se administran en forma grupal y los alumnos los contestan por sí mismos– y están compuestos fundamentalmente por reactivos de opción múltiple, salvo en el caso de la evaluación de las habilidades para la escritura que ha requerido el uso de preguntas de respuesta corta y tipo ensayo. En el caso de la educación preescolar la evaluación realizada en 2007 exigió el diseño de pruebas que no podían presuponer el dominio de la lectura y la escritura por parte de los alumnos evaluados y que requerían, por la edad de éstos y su nivel de maduración, de situaciones de evaluación más cercanas a su experiencia cotidiana. Además, el enfoque centrado en competencias del plan de estudios vigente al momento de la evaluación le dio también un carácter especial al Excale correspondiente (INEE, 2008).

Los Excale, como ya se ha dicho, son instrumentos de evaluación alineados al currículum nacional. En el caso de primaria y secundaria los exámenes fueron diseñados de acuerdo a los planes y programas de estudio vigentes en el país desde 1993 –en el ciclo escolar 2006-2007 se generalizó en México, después de un año de prueba, la Reforma de la Educación Secundaria y, actualmente, está en fase de pilotaje un nuevo currículum para la educación primaria–. Los Excale de preescolar por su parte, sí pudieron tomar como referente el nuevo programa de educación preescolar, en implementación desde el ciclo escolar 2004-2005.

TABLA 4. EJES DEL CURRÍCULUM DE MATEMÁTICAS POR NIVEL EDUCATIVO EVALUADOS A TRAVÉS DE LOS EXCALE

Nivel educativo	Ejes curriculares evaluados		
<i>Preescolar (campos formativos)</i>	1. Número		
	2. Forma, espacio y medida		
<i>Primaria (ejes temáticos)</i>		3°	6°
	1. Los números, sus relaciones y sus operaciones	X	X
	2. Medición	X	X
	3. Geometría	X	X
	4. Tratamiento de la información	X	X
	5. Predicción y azar		X
<i>Secundaria (ejes temáticos)</i>	6. Procesos de cambio		X
	1. Aritmética		
	2. Álgebra		
	3. Geometría		
	4. Presentación y tratamiento de la información		
	5. Probabilidad		

FUENTE: Backhoff y otros (2006, 2007 y 2008).

La tabla 4 muestra los grandes ejes de organización del currículum nacional de Matemáticas para cada uno de los niveles que integran la educación básica. En el caso de la educación primaria se identifican los ejes considerados en las evaluaciones de 3° y 6° dado que dos de ellos se estudian propiamente hasta el 4° grado y por consiguiente, no fueron considerados en la evaluación de 3°. Desde luego, cada uno de los campos formativos o ejes temáticos se desagregan en conjuntos mucho más finos de contenidos que constituyeron los referentes específicos para el diseño de las pruebas (véase en Backhoff y otros (2006), Backhoff y otros (2007) y Backhoff y otros (2008) información detallada sobre las especificaciones y contenido de los reactivos que integraron los exámenes).

Todos los exámenes se califican en una escala 200-800 con una media centrada en 500 puntos y una desviación estándar de 100. Sin embargo, esto no significa que las pruebas correspondientes a distintos niveles educativos, grados o asignaturas sean equiparables entre sí dado que hasta el momento los niveles de logro han sido establecidos para cada prueba en particular.

4. LOS PRINCIPALES RESULTADOS DE LAS EVALUACIONES

Los informes que han dado cuenta de los resultados de las evaluaciones realizadas entre 2005 y 2007 en la educación básica muestran en forma pormenorizada distintas clases de información, fundamentalmente por lo que toca a las métricas utilizadas para reportar los resultados, las medias poblacionales y el porcentaje de alumnos en los niveles de logro educativo, a escala nacional y desagregados según estrato escolar y entidad federativa cuando ello es posible, pero también de acuerdo a otras variables significativas como el sexo y la edad. De igual forma, es posible formarse una buena idea del desempeño de los alumnos ante los diferentes contenidos específicos que evalúan las pruebas. Las métricas utilizadas (particularmente el porcentaje de alumnos que logran los niveles de desempeño establecidos) satisfacen en mayor grado las expectativas sociales en materia de información derivada de las evaluaciones y pueden tener mayor impacto sobre los docentes, al darles “una señal acerca de qué es lo que se espera que todos los alumnos sepan y sean capaces de hacer al cabo de cada grado o ciclo de enseñanza” (Ravela, 2002) y, qué es lo que saben y son capaces de hacer a la luz de los resultados efectivamente obtenidos.

Así, en el contexto de este artículo son de particular interés los resultados por nivel de logro y estrato escolar ya que son más informativos que los promedios y dejan ver en forma más clara los desafíos globales que enfrenta hoy en día el sistema educativo mexicano para lograr una educación más eficaz. Enseguida se presentan estos resultados de acuerdo a la secuencia en que se cursan los niveles de la educación básica.

La evaluación de la calidad de la educación preescolar muestra que en el caso de las Matemáticas, a escala nacional, la mitad de los alumnos que estaban terminando ese nivel educativo se ubicaba en el nivel *básico* (véase tabla 3 y en el anexo 1 la definición de los niveles de logro para preescolar). En términos generales este nivel, como ya se ha dicho, indica un dominio elemental de los conocimientos, habilidades y destrezas escolares necesarios para progresar satisfactoriamente en la materia. Habiendo sido la evaluación en preescolar la más reciente de las referidas en este artículo, el reducido porcentaje de alumnos que para el país en su conjunto quedó ubicado en el nivel *por debajo del básico* fue un hallazgo alentador, dado que —como se muestra más adelante— en los niveles educativos evaluados previamente,

la cantidad de alumnos que quedó en aquel (especialmente en 3° de primaria y en secundaria) resultó preocupante.

TABLA 5. PORCENTAJE DE ALUMNOS DE PREESCOLAR POR ESTRATO ESCOLAR Y NIVEL DE LOGRO. PENSAMIENTO MATEMÁTICO

Estrato escolar	Nivel de logro escolar				Puntaje promedio
	Debajo del básico	Básico	Intermedio	Avanzado	
Nacional	9	49	27	15	500
Cursos comunitarios	22	56	16	6	442
Rural público	16	58	19	6	457
Urbano público	8	51	28	13	499
Privado	1	27	36	36	572

FUENTE: Backhoff y otros (2008).

El análisis de los resultados por estrato escolar da cuenta de las inequidades al interior del sistema educativo. En preescolar, al igual que en el resto de los niveles educativos, la distribución de los alumnos en los niveles de logro escolar revela grandes diferencias entre los estratos; en efecto, ya sea que se agrupen de un modo o de otro los porcentajes (de básico a avanzado, sólo intermedio y avanzado considerando que éstos son los niveles deseables) para establecer énfasis distintos, lo cierto es que la distancia entre los resultados de alumnos que acuden a escuelas ubicadas en uno u otro estrato llega a ser tan grande como la que se observa al comparar las distribuciones de los preescolares comunitarios y privados, o el puntaje promedio en cada uno de los dos estratos (distante por más de una desviación estándar).

Como se señaló en el primer apartado de este artículo, los estratos considerados en las evaluaciones de la calidad del sistema educativo dan cuenta no sólo de modelos educativos diferentes –y de ofertas escolares distintas, en un sentido más amplio–, sino también de condiciones de contexto profundamente contrastantes, lo que, sin duda, tiene un fuerte impacto en los resultados de aprendizaje revelados por las evaluaciones realizadas. En realidad, lo que hace especialmente grave este panorama es que, más allá de la dificultad para deslindar con claridad el efecto de los factores del entorno y la escuela, la gran desigualdad de los resultados escolares se explica, por la conjunción de ambos. Los alumnos que viven en circunstancias más desfavorables en el hogar son atendidos en escuelas de mayores carencias. En vez de que la escuela compense, en alguna medida, las desventajas de los alumnos de medios desfavorecidos, contribuye más bien a agudizarlas (INEE, 2007a).

Interesa pues subrayar que desde el preescolar se establecen brechas en el logro escolar y que también desde la base del trayecto educativo casi seis de cada diez niños a nivel nacional, logran apenas el aprendizaje básico, lo que significa que ciertos conocimientos o destrezas que de acuerdo al currículum del preescolar deberían lograr todos los alumnos, no están siendo alcanzados, sobre todo en los medios socioculturales donde más importante sería que la escuela compensara las desigualdades de origen.

En el caso de la educación primaria –nivel educativo que comprende el doble de grados que los que integran el preescolar y la educación secundaria- en 2005 se evaluó una muestra de alumnos del grado terminal (6°) y al año siguiente una de tercer grado. La evaluación de este último grado mostró, a nivel nacional, uno de los porcentajes más altos de alumnos en el nivel *por debajo del básico* (40%). Aunque en términos generales el análisis por estratos puso también de manifiesto el contrastante desempeño de los alumnos del medio urbano o rural y del sector de sostenimiento público o privado (véase tabla 6), el

hallazgo más sorprendente fue el notable desempeño relativo de los alumnos de cursos comunitarios (el segundo en orden descendente, sólo por debajo del estrato privado).

TABLA 6. PORCENTAJE DE ALUMNOS DE 3º DE PRIMARIA POR ESTRATO ESCOLAR Y NIVEL DE LOGRO. MATEMÁTICAS

Estrato escolar	Nivel de logro escolar				Puntaje promedio
	Debajo del básico	Básico	Intermedio	Avanzado	
Nacional	40	20	31	9	500
Educación indígena	72	14	12	2	422
Rural público	52	20	24	4	469
Urbano público	36	21	34	9	507
Cursos comunitarios	35	17	36	12	513
Privado	13	15	45	28	578

FUENTE: Backhoff y otros (2007).

El desempeño de los alumnos de tercer grado de cursos comunitarios no concuerda con las pautas identificadas en las evaluaciones de todos los demás grados y niveles educativos que incluyen este modelo educativo (preescolar y primaria), que ubican a este estrato en el último lugar o sólo por encima de las escuelas indígenas cuando este estrato ha sido incluido en la evaluación. La singularidad de la situación dio lugar a que se analizaran por separado los resultados de los cursos comunitarios y a que se formularan hipótesis para explicar estos hallazgos. Dos explicaciones tentativas –para las que no hubo evidencias que las sustentaran– se refirieron a posibles problemas técnicos en la aplicación de los Excale; la tercera hipótesis, que requiere todavía de más evidencia empírica a través del allegamiento de información adicional a la que hasta el momento se ha reunido y analizado (Backhoff y otros, 2007; INEE, 2007a), es que los resultados revelan realmente un mejor desempeño de los alumnos de tercer grado que asisten a los cursos comunitarios.

En cualquier caso, y salvo este resultado, la desigual distribución de los aprendizajes reportada antes para el caso de la educación preescolar se sostiene –como sucede en forma sistemática en realidad, a lo largo de toda la educación básica– en la evaluación del aprendizaje en 3º de primaria (véase en el anexo 1 la definición de los niveles de logro para este grado).

Los resultados de la evaluación realizada en 2005 en el grado terminal de la educación primaria muestran que la distribución de los alumnos de distintos estratos en los niveles de logro escolar establecidos sigue pautas consistentes con el conocido y fuerte influjo de las condiciones de origen social en los resultados de aprendizaje (véase tabla 7), de suerte que son los alumnos de escuelas indígenas los que obtienen un menor puntaje promedio y quedan ubicados en mayor porcentaje (muy cercano a la mitad de los alumnos de esta modalidad) en el nivel *por debajo del básico*, mientras que los de escuelas privadas vuelven a alcanzar las mejores posiciones.

TABLA 7. PORCENTAJE DE ALUMNOS DE 6° DE PRIMARIA POR ESTRATO ESCOLAR Y NIVEL DE LOGRO. MATEMÁTICAS

Estrato escolar	Nivel de logro escolar				Puntaje promedio
	Debajo del básico	Básico	Intermedio	Avanzado	
Nacional	17	52	24	7	500
Educación indígena	43	49	7	1	424
Cursos comunitarios	28	58	13	1	457
Rural público	24	57	16	3	471
Urbano público	14	53	26	7	510
Privado	3	31	42	24	589

FUENTE: Backhoff y otros (2006).

En el conjunto del país poco más de la mitad de los alumnos de 6° grado se ubica en el nivel de logro escolar *básico*, 17% en el más bajo, y casi una tercera parte en los niveles *intermedio* y *avanzado*. El bajo porcentaje de alumnos de 6° que quedan en el nivel *básico* llama la atención si se compara con el dato análogo para 3° de primaria y suscita algunas preguntas: ¿este resultado es atribuible a las pruebas aplicadas o al propio currículum nacional? ¿La evaluación de 3° fue particularmente difícil? ¿Los niveles de logro establecidos para cada prueba/grado suponen distintos niveles de exigencia? ¿El currículum adolece de problemas de graduación de los contenidos? El propio INEE adelanta algunas hipótesis, relacionadas efectivamente con esos dos ámbitos: el del currículum y su enseñanza por un lado, y el de la propia evaluación, sus avances y pendientes, por otro lado.

La evaluación en secundaria arroja resultados coincidentes en lo esencial con lo ya señalado para el preescolar y la primaria. Si bien, en este caso –como fue advertido en el apartado 1 de este artículo– los estratos escolares están definidos fundamentalmente en términos de los modelos educativos que se implementan en ellos, las variables de contexto que subyacen a las categorías definidas para preescolar y primaria están también presentes en la educación secundaria dada la población escolar que fundamentalmente se atiende en cada estrato (el tipo de sostenimiento y la ubicación urbana o rural). De hecho, es plausible suponer que en su mayoría, los egresados de las escuelas primarias urbano-públicas ingresan a secundarias generales o técnicas, los de escuelas indígenas y cursos comunitarios a telesecundarias y los de primarias privadas continúan en secundarias de este sector.

TABLA 8. PORCENTAJE DE ALUMNOS DE 3° DE SECUNDARIA POR ESTRATO ESCOLAR Y NIVEL DE LOGRO. MATEMÁTICAS

Estrato escolar	Nivel de logro escolar				Puntaje promedio
	Debajo del básico	Básico	Intermedio	Avanzado	
Nacional	51	30	18	1	500
Telesecundaria	62	26	11	1	473
Técnica	52	30	17	1	497
General	50	31	18	1	500
Privado	24	31	38	7	574

FUENTE: Backhoff y otros (2006).

Nuevamente el porcentaje de alumnos que a nivel nacional se ubica en la categoría *por debajo del básico* es elevado (más que en el tercer grado de primaria). La prueba de Matemáticas para 3° de secundaria –aplicada en 2005 al igual que la de 6° de primaria–, es, de acuerdo a sus propiedades psicométricas, medianamente adecuada para la población de este grado escolar. En el informe correspondiente se señala que siendo los Excale pruebas alineadas al currículum, esto puede deberse a dos razones: i) que

los reactivos fueron elaborados con un mayor nivel de dificultad que el establecido en los planes y programas de estudio; ii) que los alumnos no adquieren los conocimientos y desarrollan en grado adecuado las habilidades matemáticas señaladas en dichos planes y programas. Los autores consideran que por la metodología empleada –que supone la validación de los Excale por docentes en ejercicio– la segunda explicación es más plausible que la primera (Backhoff y otros, 2006).

Es indudable que se requiere un mayor y más rico esfuerzo de evaluación e investigación para responder con más fundamento las preguntas que las propias evaluaciones y sus resultados van suscitando. Lo que está claramente sustentado con el conocimiento que ya se ha generado hasta el momento por las evaluaciones nacionales realizadas en el país es que hay sectores amplios de la población escolar – particularmente amplios en algunos grados escolares como 3º de primaria y 3º de secundaria– que no están logrando aún el nivel más fundamental, básico o elemental en su aprendizaje y que esta situación compromete tanto su propia trayectoria escolar como su inserción plena, productiva y en la vida social.

También es claro que el sistema educativo, al menos desde la perspectiva de sus logros es inequitativo, a juzgar por las enormes brechas en el desempeño escolar de los alumnos atendidos por los distintos tipos de servicio educativo que hoy en día constituyen la oferta de educación básica en el país.

5. REFLEXIONES FINALES

Los resultados de evaluación de que se da cuenta en este artículo se refieren exclusivamente al ámbito de las Matemáticas. Las razones para hacer esta selección mencionadas en la Introducción son, la mayor importancia relativa de la escuela para explicar las diferencias de aprendizaje en este campo; y, por otro lado, la menor atención de que ha sido objeto –al menos en comparación con el Español y las habilidades comunicativas– por las políticas educativas. Ahora bien, con matices y rasgos particulares en cada caso, lo cierto es que el grueso de los resultados –el desigual desempeño de los alumnos de distintos estratos escolares, la proporción menor o mayor pero siempre significativa de alumnos cuyo nivel de logro es básico o inferior–, está presente en todas las evaluaciones realizadas. El menor desempeño de los alumnos de escuelas indígenas y cursos comunitarios o de telesecundarias y la enorme brecha entre éste y el que muestran los de instituciones privadas es una constante ya sea que se evalúen resultados en Matemáticas, Español o cualquier otra asignatura o área de desarrollo.

Los rezagos que enfrenta el sistema educativo mexicano en la calidad de la educación básica que ofrece son pues generalizados y revelan problemas tanto de eficacia como de equidad. Sin embargo, es pertinente insistir en la importancia relativa del área de Matemáticas y las habilidades que ésta engloba. Fundamentalmente para llamar la atención sobre la necesidad de prestar mayor atención a la formación de los educandos en ella a través de estrategias integrales, que impliquen tanto al currículo nacional en su enfoque de enseñanza, propósitos formativos, contenidos y criterios de desempeño y evaluación, como otros aspectos de importancia no menor, como la formación y actualización docentes, la provisión de materiales didácticos y, en general, de mejores condiciones de enseñanza y aprendizaje.

No sobra señalar, que los aprendizajes que logran los alumnos en los grados de egreso de la educación primaria y secundaria (6º y 3º respectivamente), son de particular importancia porque para cierto número de niños y jóvenes siguen representando hoy en día la escolaridad máxima que alcanzarán, muestran los rezagos en la formación identificados en la evaluación realizada en 2005.

Por otro lado, y como también se apuntó en este artículo, los desafíos para las propias evaluaciones son claros. Los estudios realizados en los tramos que conforman la educación básica (y dentro de éstos, en grados clave como los terminales) forman parte de un programa de evaluación del aprendizaje que intenta mostrar, a lo largo del tiempo, las tendencias de la calidad de la educación que se ofrece en México; un esfuerzo por analizar los resultados desde esta perspectiva puede verse en una de las publicaciones más recientes del INEE (2008). Sin embargo, lo cierto es que todavía tienen que desplegarse acciones adicionales para contar con instrumentos de evaluación equiparados tanto vertical como horizontalmente, que permitan comparar con mayor fundamento el desempeño de un mismo grupo poblacional entre distintas áreas de aprendizaje, o bien, el de grupos diferentes con distinta escolaridad. Sólo entonces se podrá afirmar con mayor seguridad la dirección de las tendencias de aprendizaje de los alumnos atendidos por el sistema educativo mexicano.

REFERENCIAS BIBLIOGRÁFICAS

- Banco Mundial (2008). *Educación Media en el Paraguay. Logros, opciones y desafíos*. Informe N° 42665-PY. Washington: Banco Mundial.
- Backhoff, Eduardo y otros (2008). *El aprendizaje en tercero de preescolar en México. Lenguaje y comunicación. Pensamiento matemático*. México: INEE.
- Backhoff, Eduardo y otros (2007). *El aprendizaje en tercero de primaria en México: Español, Matemáticas, Ciencias Naturales y Ciencias Sociales*. México: INEE.
- Backhoff, Eduardo y otros (2006). *El aprendizaje del Español y las Matemáticas en la educación básica en México*. México: INEE.
- DGEI (2007). *Estadística total de educación indígena inicial y básica. Inicio de ciclo escolar 2007-2008*. Dirección General de Educación Indígena, México: SEP. (http://basica.sep.gob.mx/dgei/pdf/eventos/prontuario_estadistico_07_08.pdf).
- INEE (2005). *La calidad de la educación básica en México. Informe Anual 2005*. México: INEE.
- INEE (2007a). *La educación para poblaciones en contextos vulnerables. Informe Anual 2007*. México: INEE.
- INEE (2007b). *Panorama Educativo de México 2007. Indicadores del Sistema Educativo Nacional*. México: INEE.
- INEE (2008). *¿Avanza o retrocede la calidad educativa? Tendencias y perspectivas de la educación básica en México. Informe Anual 2008*. México: INEE.
- Jornet, Jesús y Eduardo Backhoff (2006). *Manual técnico para el establecimiento de niveles de competencia*. México: INEE.
- Ravela, Pedro (2002). *¿Cómo presentan sus resultados los sistemas nacionales de evaluación educativa en América Latina?* Santiago: PREAL.
- Ruiz, Guadalupe (2003). Proyectos de evaluación de la educación básica en la década 1990-2000. En: Zorrilla, Margarita (coord.). *La evaluación de la educación básica en México 1990-2000. Una mirada a contraluz*. México: Universidad Autónoma de Aguascalientes.

ANEXO 1

DEFINICIÓN DE LOS NIVELES DE LOGRO DE LOS EXCALE DE MATEMÁTICAS

PREESCOLAR

Avanzado
<p>En relación con el aspecto de Número, las alumnas y los alumnos de este nivel son capaces reutilizar números para representar cantidades hasta veinte.</p> <p>En relación con el aspecto de Forma, espacio y medida, las alumnas y los alumnos de este nivel son capaces de trazar trayectos a partir de puntos de referencia espaciales que incluyen direccionalidad (desde, hacia, hasta) y ubicar los días de la semana a partir de las actividades que realizan.</p>
Medio
<p>En relación con el aspecto de Número, las alumnas y los alumnos de este nivel son capaces de utilizar números para representar cantidades menores de 13; identificar y representar numerales hasta treinta; emplear los números para identificar hasta la cuarta posición de un elemento en una serie ordenada; resolver problemas que implican combinar el valor de monedas de cincuenta centavos, uno y cinco pesos; identificar patrones diseñados en función de un criterio de repetición de dos o tres elementos; y resolver problemas que implican agregar, igualar, comparar, quitar o repartir cantidades de una o más colecciones de hasta nueve objetos con o sin el apoyo de la representación gráfica.</p> <p>En relación con el aspecto de Forma, espacio y medida, las alumnas y los alumnos de este nivel son capaces de reconocer figuras geométricas que comparten atributos; identificar figuras semejantes a una de muestra; identificar los cambios que ocurren en una figura geométrica al combinarla con otras; resolver problemas que implican medir longitudes tomando en cuenta una unidad de medida no convencional; identificar el orden de puntos de referencia espacial en un trayecto representado gráficamente; y establecer e identificar la secuencia de imágenes que representan las relaciones temporales antes-después-al final.</p>
Básico
<p>En relación con el aspecto de Número, las alumnas y los alumnos de este nivel son capaces de utilizar números para representar cantidades menores de siete; contar colecciones de objetos, hasta treinta; comparar colecciones de objetos y establecer relaciones de igualdad y desigualdad; registrar la cantidad de elementos en tablas y gráficas, con ayudas como categorías establecidas y ejemplos; identificar en tablas o gráficas la colección en la que hay más o hay menos elementos; identificar la colección faltante en una serie de colecciones con patrón de crecimientos $n + 1$; y construir la colección que sigue en una serie de colecciones con patrón de crecimiento $n + 1$.</p> <p>En relación con el aspecto de Forma, espacio y medida, las alumnas y los alumnos de este nivel son capaces de identificar semejanzas entre un cuerpo geométrico y un objeto del entorno; identificar posiciones de objetos respecto a otros objetos, en una representación gráfica; comparar de manera perceptual la longitud de objetos: más corto que..., más largo que...; resolver problemas que impliquen estimar longitudes; y distinguir el instrumento apropiado para medir el peso.</p>
Por debajo del básico
<p>En relación con el aspecto de Número, las alumnas y los alumnos de este nivel son capaces de decir la serie numérica empezando por el uno, y hasta el que saben (máximo treinta); enumerar de manera oral objetos desde el uno y hasta el que saben (máximo veinte), siguiendo el orden de la serie numérica; identificar los usos de los números en situaciones cotidianas; y escribir en orden un tramo de la serie numérica convencional (máximo treinta) o números que se les dictan, sin embargo no logran usarlos para designar la cantidad de una colección.</p> <p>En relación con el aspecto de Forma, espacio y medida, las alumnas y los alumnos de este nivel son capaces de reconocer un objeto que se les presenta gráficamente, desde distintos puntos espaciales; e identificar de manera perceptual propiedades medibles contrastantes de los objetos (lleno-vacío, alto-bajo, largo-corto...).</p>

FUENTE: Backhoff y otros (2008).

3° de Primaria

Avanzado
Los alumnos de este nivel completan series numéricas con una constante aditiva; identifican la operación que resuelve un problema que puede ser resuelto a partir de una operación dada; resuelven problemas que implican una multiplicación con factores de dos por tres cifras con el algoritmo convencional; resuelven problemas aditivos sencillos con más de una operación, y comparan y establecen equivalencias entre medios, cuartos y octavos para expresar el resultado de problemas de reparto y medida. También reconocen el número de caras de prismas rectangulares; ubican objetos que cumplan con dos condiciones, utilizando expresiones tales como derecha-izquierda, arriba-abajo, delante de-atrás de, cuando el objeto de referencia coincide con la lateralidad del alumno; además, resuelven problemas que impliquen estimar el peso de objetos utilizando unidades de referencia no convencionales y problemas que impliquen comparar el peso de objetos utilizando $\frac{1}{2}$ kg y $\frac{1}{4}$ kg; así como problemas que impliquen el uso del reloj de manecillas (horas y minutos) y el calendario (meses, semanas y días); de igual forma, interpretan pictogramas donde el dibujo representa más de un elemento.
Medio
Los alumnos de este nivel utilizan el valor relativo de los números para construir y descomponer cantidades; leen y escriben números de cuatro cifras con cero intermedio; ordenan y comparan números de cuatro cifras; calculan con transformación mediante el algoritmo convencional; calculan multiplicaciones sencillas con factor de una cifra; identifican fracciones (medios, cuartos y octavos) en contextos de medida y reparto, expresándolos con símbolos numéricos. Asimismo resuelven problemas aditivos con la incógnita en el estado intermedio y problemas multiplicativos sencillos, utilizando procedimientos no convencionales. También reconocen la forma de las caras de cuerpos geométricos; identifican figuras básicas con relación a las características geométricas y número de sus lados; ubican objetos utilizando una condición de expresiones, tales como derecha-izquierda, arriba-abajo, delante de-atrás de, cuando el objeto de referencia no coincide con la lateralidad del alumno; además, comparan, estiman y miden longitudes utilizando el centímetro; tienen la noción del metro; resuelven problemas que impliquen la comparación del peso de objetos, utilizando el kilogramo como unidad de referencia; resuelven problemas utilizando la hora y el minuto como unidades de tiempo, e identifican lo que se puede o no contestar a partir de la información proporcionada por una gráfica o una imagen.
Básico
Los alumnos de este nivel identifican el antecesor y el sucesor de un número de tres cifras; ordenan y comparan números de tres cifras; leen y escriben números de cuatro cifras sin cero intermedio; calculan sumas con transformación y restas sin transformación mediante el algoritmo convencional; además resuelven problemas aditivos hasta de tres cifras cuya solución implica una suma con incógnita en el estado final, y realizan repartos equitativos con ayuda gráfica, utilizando modelos continuos y discretos. Por otro lado, identifican figuras básicas (círculos, triángulos, cuadrados y rectángulos) en composiciones; comparan y miden longitudes utilizando el centímetro, resuelven problemas utilizando la hora y el minuto como unidades de tiempo, y leen e interpretan pictogramas sencillos y gráficas de barras.
Por debajo del básico
Los alumnos de este nivel pueden leer y escribir números de tres cifras; calculan sumas sin transformación mediante el algoritmo convencional y resuelven problemas aditivos (sencillos) ¹ con números naturales, utilizando procedimientos informales. También identifican objetos cotidianos desde diferentes perspectivas; comparan longitudes y superficies de manera perceptual y utilizando unidades no convencionales, y responden preguntas referidas a la frecuencia partir de información contenida en tablas.

FUENTE: Backhoff y otros (2007).

¹ Llamamos "problemas sencillos" a aquellos que utilizan contextos conocidos (dinero, medida), números de tres cifras y en los que la incógnita se encuentra en el estado final.

6° de Primaria

Avanzado
Los alumnos de este nivel resuelven problemas que impliquen varias operaciones con números naturales, fraccionarios y decimales. También tienen nociones depuradas de conceptos tales como: perímetro, área y volumen; también interpretan la representación plana de un cuerpo geométrico y el desarrollo plano de una Figura; asimismo, describen trayectos en planos y mapas; pueden además realizar conversiones de unidades de medida. También interpretan información contenida en gráficas y tablas, y resuelven problemas de probabilidad que impliquen un análisis combinatorio; aplican las propiedades de la proporcionalidad.
Medio
Los alumnos de este nivel leen, comparan y ordenan números decimales y fraccionarios, y resuelven con ellos problemas sencillos de suma y resta; además resuelven problemas con números naturales que impliquen dos o tres operaciones. Igualmente clasifican Figuras con base en sus propiedades geométricas; también calculan áreas mediante el uso de fórmulas, y calculan volúmenes de Figuras mediante el conteo de unidades cúbicas; identifican puntos de croquis, planos y mapas, así como puntos en el primer cuadrante de un plano cartesiano. Asimismo, interpretan información contenida en gráficas y tablas que contienen datos; resuelven problemas sencillos de probabilidad que no impliquen realizar un análisis combinatorio; y resuelven problemas de proporcionalidad.
Básico
Los alumnos de este nivel leen, ordenan y comparan números naturales; además resuelven problemas sencillos con números naturales, decimales y fraccionarios que impliquen una operación en contextos conocidos. Adicionalmente, calculan perímetros y áreas de triángulos y cuadriláteros dentro de una retícula. Asimismo, interpretan información contenida en gráficas y tablas sencillas.
Por debajo del básico
Los alumnos de este nivel resuelven problemas con una operación que implique sumas o restas con números de hasta cuatro cifras; además comparan decimales con el mismo número de cifras. Asimismo, calculan el promedio de números naturales en contextos conocidos.

FUENTE: Backhoff y otros (2007)

3° de Secundaria

Avanzado
Resuelven problemas que implican potenciación y radicación con números naturales o decimales. Dividen polinomios y modelan problemas mediante ecuaciones de segundo grado. Establecen relaciones entre todo tipo de representaciones tabulares, gráficas y algebraicas de una función lineal o cuadrática. Factorizan polinomios. También resuelven problemas que implican el cálculo del área del círculo, así como del área lateral y volumen de cuerpos geométricos; también utilizan propiedades o teoremas sencillos para resolver problemas geométricos o de medición (por ejemplo, las razones trigonométricas y el teorema de Pitágoras). Realizan transformaciones o movimientos en el plano (simetría, rotación, traslación) en la resolución de problemas de construcción, de medida o de escala. Finalmente, resuelven problemas de probabilidad aplicando la regla de la suma o del producto.
Medio
Los alumnos de este nivel resuelven problemas en los que se utilizan varias operaciones con números decimales o fraccionarios así como la raíz cuadrada con números naturales. Resuelven problemas de proporcionalidad con valor unitario no entero o de reparto proporcional, incluyendo porcentajes. Resuelven problemas que implican usar la jerarquía de operaciones. Asimismo, multiplican polinomios y modelan situaciones mediante una función lineal o cuadrática. Modelan también situaciones que implican sistemas de ecuaciones con dos incógnitas. Establecen relaciones entre dos cualesquiera de las formas de representación de una función lineal o cuadrática. Establecen relaciones entre una tabla de valores y su expresión algebraica. Asimismo, resuelven problemas que implican el cálculo del perímetro del círculo, utilizan las propiedades de lugares geométricos (alturas, bisectrices, mediatrices) en la resolución de problemas de construcción, de medida o de escala. Finalmente, interpretan información contenida en tablas o gráficas de distintos tipos. Resuelven problemas de conteo y determinan la probabilidad de eventos mutuamente excluyentes.
Básico
Los alumnos de este nivel resuelven problemas que implican dos o más operaciones (adición, sustracción, multiplicación, división y potenciación) con números naturales y enteros. Identifican situaciones de proporcionalidad. Suman y restan polinomios, y resuelven ecuaciones de primer grado con una incógnita, así como sistemas de ecuaciones con dos incógnitas. Igualmente, utilizan las propiedades de las Figuras (por ejemplo: ángulos, lados y diagonales) en la resolución de problemas de medición. También calculan el perímetro y área de figuras básicas (triángulos, cuadriláteros y polígonos regulares) y el volumen de cuerpos geométricos.
Por debajo del básico
Los alumnos de este nivel resuelven problemas que implican una adición o una sustracción con números naturales, enteros, decimales o fraccionarios, así como problemas de multiplicación y división con números naturales. Asimismo, establecen relaciones entre una tabla de valores y su gráfica, en funciones lineales o cuadráticas. Adicionalmente, identifican Figuras o cuerpos geométricos a partir de sus elementos (lados, caras, etcétera). Finalmente, estiman y comparan la probabilidad de eventos simples.

FUENTE: Backhoff y otros (2006).

HONDURAS 2008: LA EVALUACIÓN DE LOS APRENDIZAJES VISLUMBRA UNA LUZ DE ESPERANZA EN EDUCACIÓN

Russbel Hernández

Revista Iberoamericana de Evaluación Educativa 2009 - Volumen 2, Número 1

<http://www.rinace.net/riee/numeros/vol2-num1/art5.pdf>

Fecha de recepción: 13 de marzo de 2009

Fecha de comunicación de dictamen: 12 de abril de 2009

Fecha de aceptación: 22 de abril de 2009

En Honduras, durante la última década los resultados de la evaluación externa de los aprendizajes en Español y Matemáticas habían estado estancados en valores muy bajos cercanos al 40%, como promedio nacional. Sin embargo, la introducción de un nuevo currículo de educación básica, a partir del 2004, mediante un sistema de materiales educativos alineados a dicho currículo y que sirven de apoyo a la labor docente en el aula (incluyendo Libros de Texto, Guías didácticas para el docente, Estándares Educativos, Programaciones Educativas Mensuales, Estándares de Desempeño, Pruebas Formativas Mensuales y Cuadernos de Trabajo para alumnos) parece haber cambiado esa tendencia. El desarrollo e implementación de este sistema de materiales educativos ha sido posible con el apoyo de la cooperación internacional como parte de la iniciativa Education for All (EFA) que cuenta con apoyo directo de la Cooperación Canadiense, Española, Europea, Sueca y Alemana, y de proyectos específicos como PROMETAM de JICA y MIDEH de USAID, por lo que puede considerarse como evidencia de una intervención exitosa en un sistema educativo nacional.

Enmarcados en el nuevo Curriculum Nacional Básico, es que se han elaborado pruebas estandarizadas en Matemáticas y Español para los nueve (9) grados de la educación básica. En el año 2007, se aplicaron pruebas a 1ero, 3ero y 6to grado en las escuelas públicas y en el año 2008 se aplicaron las pruebas de 1ero a 6to grado en una muestra probabilística a nivel nacional en los cuales se consideró la estratificación según las regiones o 18 departamentos del país, según la ubicación geográfica (urbano-rural) y según el tipo de administración (pública y privada). Además se validaron las pruebas de 7mo a 9no grado también en una muestra probabilística con representatividad a nivel nacional.

La aplicación de las pruebas estandarizadas se llevó a cabo bajo la coordinación de la Dirección General de Evaluación de la Calidad de la Educación (DIGECE) y ejecutado por el Proyecto Mejorando el Impacto al Desempeño Estudiantil en Honduras (MIDEH) que es financiado por USAID para apoyar a la Secretaría de Educación de Honduras.

El presente artículo, presenta los resultados de la evaluación externa y estandarizada en Matemáticas y Español correspondiente al año 2008 en todos los seis (6) grados iniciales de la educación básica hondureña. Sin embargo, en los dos primeros hallazgos se hace uso de los hallazgos o resultados encontrados en 1ero, 3ero y 6to grado de las escuelas públicas, ya en el año 2007 solamente a este tipo de centros educativos se evaluó en una muestra probabilística a nivel nacional.

1. MARCO CONTEXTUAL

“Es urgente mejorar la calidad educativa” planteó un informe de Progreso Educativo en Honduras (PREAL, 2006:13), destacando los escasos avances que en esta materia ha tenido el país, pese a su permanente abordaje en la agenda social y política nacional durante las últimas dos décadas. Una breve panorámica de la calidad educativa siguiendo la definición propuesta por Schmelkes (*et al.*, 1989), que

considera cuatro grandes dimensiones Eficiencia, Eficacia, Equidad y Relevancia, evidencia el porqué de la urgencia.

La dimensión **Eficiencia**, referida al hecho que debe procurarse el logro de los objetivos educacionales con el mínimo de insumos requeridos para ello, siendo la reprobación, repitencia, rezago, deserción, así como la relación entre los niveles de inversión por alumno y los niveles de aprendizaje alcanzados, son algunos de los indicadores que evidencian la crítica situación de la educación hondureña en esta área. En el año 2008 se han logrado algunos avances significativos en el I y II Ciclo de la Educación Básica, tal como se observa en el siguiente cuadro.

CUADRO 1. INDICADORES DE EFICIENCIA POR NIVEL EDUCATIVO. HONDURAS. 2008

Nivel	Reprobación	Repitencia	Deserción
I-II Ciclo Ed. Básica	4.0%	5.0%	0.88%
III Ciclo Ed. Básica	4.0%		13.51%
Nivel Medio	14.8%		7.46%

FUENTE: Secretaría de Educación. Unidad de Planificación y Evaluación de la Gestión (UPEG)-Dirección General de Evaluación de la Calidad de la Educación (DIGECE). 2009

En términos de **Equidad**, referida a la atención de los educandos en forma diferenciada de acuerdo con sus características individuales y a su entorno social, los indicadores de distribución de recursos de infraestructura, didácticos, humanos, la inversión por alumno, etc., evidencian claramente los problemas del sistema educativo nacional. Un indicador muy significativo en esta área es el hecho de que un docente de primaria del área urbana atiende en promedio 1.1 grados simultáneamente, mientras que su colega del área rural lo hace con 2.5 grados como media (Alas, 1999:42). La preparación de los docentes también es diferenciada entre las áreas urbano – rural con desventaja para esta última, con mayores proporciones de “empirismo” en todos los niveles.

Y en términos de **Relevancia**, referida al hecho de que los objetivos y contenidos educacionales deben considerarse en relación con su aporte para la vida en sociedad de los alumnos, como agentes sociales, como trabajadores productivos y como ciudadanos responsables, el sistema educativo también presenta indicadores críticos. En educación media se ha señalado que “existe un marcado desfase entre las principales actividades económicas de cada departamento del país y las opciones curriculares que concentran la mayor proporción de matrícula del nivel” (S.E., 2004:149).

En Primaria (I-II Ciclo de Educación Básica) en cambio se ha venido trabajando con programas de estudio aprobados en 1967 que fueron actualizados y revisados en 1987-1989, a partir de los cuales se elaboraron los Rendimientos Básicos de mayor complejidad (lo que constituyó un avance significativo en su momento). En el año 2003 se elaboró el Diseño del Currículo Nacional Básico que se distribuyó a nivel nacional en el 2005 y a partir del cual se están generando importantes materiales educativos en la actualidad.

En la dimensión de **Eficacia**, entendida como el grado de cobertura educacional y los niveles de aprendizaje alcanzados (es decir, con el grado de logro de los objetivos educativos propuestos), es claro que el sistema educativo hondureño aún necesita mejorar, tal como se evidenciaba en el año 2004.

CUADRO 2. INDICADORES DE EQUIDAD: COMPARACIÓN DE LA COBERTURA POR REGIÓN URBANO – RURAL Y NIVEL EDUCATIVO. HONDURAS. 2004

Nivel	Preescolar	Primaria (1° - 6°)	Secundaria (7° - 9°)	Media (10° - 12°)
Área Urbana	42.4%	90.5%	57.6%	35.3%
Área Rural	32.	88.6%	38.2%	8.6%

FUENTE: PREAL – FEREMA. 2005. *Informe de Progreso Educativo: Honduras*. p. 41 en base a INE (EPH 2004)

En el 2004, la cobertura por nivel también evidenciaba marcadas diferencias en detrimento del área rural (en Preescolar 42.4% contra 32.1%, en Primaria 90.5% contra 88.6%, en 7° - 9° grado 57.6% contra 38.2% y en Media 35.3% respecto a 8.6%)¹

Sin embargo, hay que destacar los avances obtenidos en los últimos 4 años, especialmente en Preescolar o Prebásica y en el I-II Ciclo de Educación Básica.

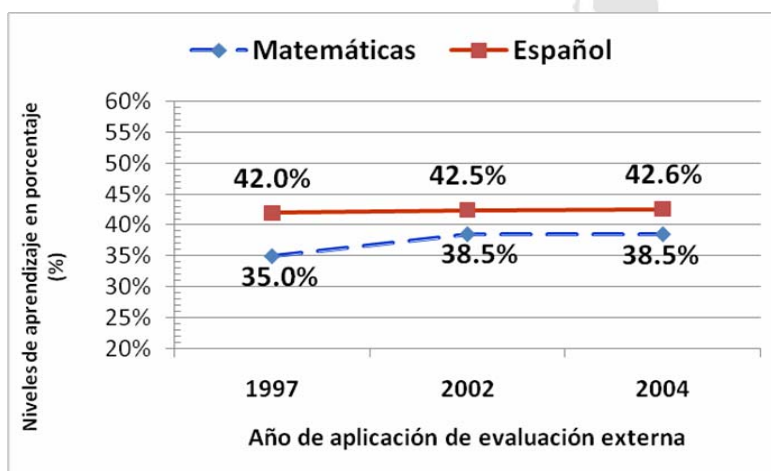
CUADRO 3. COBERTURA NETA POR NIVEL EDUCATIVO. HONDURAS. 2008

Nivel	Preescolar	I-II Ciclo E.B. (1° - 6°)	III Ciclo E.B. (7° - 9°)	Media (10° - 12°)
Porcentaje	63.7%	94%	38%	28%

FUENTE: Secretaría de Educación. Unidad de Planificación y Evaluación de la Gestión (UPEG)-Dirección General de Evaluación de la Calidad de la Educación (DIGECE). 2009

Además, los resultados de las pruebas nacionales de aprendizaje en Matemáticas, Español y Ciencias Naturales correspondientes al año 2002 han sido críticos, con medias de entre 40% y 45% (se considera un mínimo de 60% como aprobado). En ellas, menos de uno de cada cinco alumnos evaluados muestra un nivel de suficiencia ($\geq 60\%$ en las pruebas), con tasas de apenas 9.3% y 9.9% en Español de tercero y sexto grado respectivamente, 15.9% y 8.2% en Matemáticas de esos grados y 14.9% y 17.2% en Ciencias Naturales de los mismos grados.

GRÁFICO 1. RENDIMIENTO ACADÉMICO PROMEDIO EN ESPAÑOL Y MATEMÁTICAS. ESTUDIANTES DE 3° Y 6° GRADO DE CENTROS EDUCATIVOS DE ADMINISTRACIÓN PÚBLICA, HONDURAS 1997, 2002, 2004



FUENTE: Secretaría de Educación/ UMCE. 2008. Documento multicopiado.

¹ PREAL – FEREMA. 2005. *Informe de Progreso Educativo: Honduras*. p. 41 en base a INE (EPH 2004).

Y estos resultados tan bajos en las evaluaciones externas de aprendizaje mostraron una tendencia a estancarse durante casi una década (como lo ilustra el gráfico anterior), en ambas asignaturas; Español y Matemáticas. Valores que invitaban al pesimismo tanto por su significado inmediato en términos de niveles de aprendizaje alcanzados, como por la inmovilidad de los mismos que indicaban, prácticamente que no se estaba teniendo ningún avance en esta área. En este marco de resultados tan pobres y estancados, se establecieron las METAS EFA (Education for All) para el Rendimiento Académico en evaluación externa a partir del año 2005, definiéndose como meta final para el año 2015 llegar a un promedio de 1° a 6° grado de 70% en los resultados de evaluación externa de los aprendizajes en el I-II Ciclo de la Educación Básica.

Año referido	META EFA para Rendimiento Académico (evaluación externa)
2005	50%
2006	52%
2007	54%
2008	55%
2011	61%
2015	70%

FUENTE: Secretaría de Educación. 2003. Plan Todos con Educación Honduras 2003-2015 FTI-EFA.

Sin embargo, a partir del año 2004 dos nuevos factores aparecieron en el escenario educativo hondureño: Se estructuró un nuevo currículo, el Currículo Nacional Básico CNB, y se inició el desarrollo y distribución de todo un sistema de materiales educativos alineados con el CNB para apoyar la labor docente en el aula.

En relación con el primer factor, y en el marco de este deficiente contexto educativo, la Secretaría de Educación, responsable de normar los procesos del Sistema Educativo Nacional, consciente de los desafíos y las necesidades de cara al siglo XXI, inició a partir del año 2000, un proceso de transformación y reforma de la educación en Honduras orientada a que los niños, niñas y jóvenes desarrollen las competencias básicas necesarias para su adaptación a las actuales exigencias nacionales e internacionales. En este contexto, la Secretaría de Educación estructuró el Currículo Nacional Básico (2004) y su correspondiente Diseño Curricular (el DCNB). Un antecedente importante del DCNB fue la propuesta de FONAC (Foro Nacional de Convergencia) para la transformación de la educación hondureña (2000).

El segundo factor referido se desarrolló considerando que ante esta acentuada crisis de la calidad de la educación hondureña cabían plantearse varias líneas de trabajo, una de las cuales era el desarrollo y aplicación de materiales educativos para los docentes y alumnos. Dado que en educación básica, esta opción ha mostrado ser muy relevante, pues como lo ha señalado el informe de la UNESCO **"La educación encierra un tesoro"**, éste es el factor más importante en relación con el aprendizaje de los educandos (1996:135). Este hecho ha sido corroborado por numerosos estudios realizados en Latinoamérica, tal como lo han destacado Vélez y Schiefelbein en su trabajo sobre los **"Factores que afectan el Rendimiento Académico en la Educación Primaria"**.²

² En dicho documento se destaca que en 13 de 17 estudios sobre la relación entre el uso de materiales educativos y el rendimiento se ha encontrado una asociación positiva significativa (1996, p. 13).

Precisamente, desarrollando esta línea de trabajo, en la actualidad hay varias instituciones y proyectos que están colaborando con el país, para mejorar los indicadores educativos. Así, la cooperación Japonesa JICA a través del PROMETAM (Proyecto Mejoramiento de la Enseñanza Técnica en el Área de Matemática) ha elaborado textos de Matemáticas alineado con el nuevo currículo, los cuales han sido entregados en los centros educativos desde el año 2006. La Secretaría de Educación ha entregado también libros de texto de Español alineados con el DCNB, para los seis primeros grados de educación básica.

Además, el Proyecto MIDEH (Mejorando el Impacto al Desempeño Estudiantil de Honduras), está apoyando al sistema educativo hondureño a través de la elaboración de los Estándares Educativos en Matemáticas y Español, Programaciones Educativas Mensuales (que organizan los Estándares de acuerdo con planificaciones mensuales), Pruebas Formativas Mensuales (las cuales se constituyen en una innovación en materia de evaluación en el país), así como también la elaboración de los instructivos para el uso e implementación de los apoyos didácticos anteriormente mencionados, los cuales se constituyen en un gran aporte para los docentes y alumnos de la educación básica hondureña.

FIGURA 1. SISTEMA DE MATERIALES EDUCATIVOS DE APOYO A LA LABOR DOCENTE EN EL AULA



Cabe destacar que la alineación de todos los elementos pedagógicos antes mencionados con el DCNB, ofrece una visión de reforma sistémica de la educación que, de complementarse con un programa alineado de capacitación, seguimiento y supervisión de los docentes, promete grandes esperanzas de mejoras sustanciales en los logros de los alumnos y alumnas y en la calidad de la enseñanza hondureña. Al mismo tiempo, este alineamiento entre el DCNB, los Estándares Educativos, Libros de texto de Español y Matemáticas, Programaciones Educativas y Pruebas Formativas, posibilita la implementación de un sistema de **evaluación externa con Pruebas Estandarizadas**, que permitan conocer cuánto, qué, quiénes, etc., están aprendiendo los alumnos, y poder orientar la toma de decisiones correspondientes.

En ese sentido, los resultados de la evaluación externa y estandarizada de los aprendizajes realizada en el año 2008 que aquí son presentados constituyen el segundo esfuerzo de evaluación externa desde el inicio de la implementación del DCNB en los centros educativos del país (el primero fue en el año 2007), e incluyó una muestra nacional de centros educativos de los 18 departamentos del país, evaluando alumnos de primero a sexto grado en las asignaturas de Español y Matemáticas. Esta evaluación externa fue coordinada por la Dirección General de Evaluación de la Calidad de la Educación (DIGECE) y ejecutada por el equipo técnico del Proyecto Mejorando el Impacto al Desempeño Estudiantil en Honduras (MIDEH).

2. MARCO CONCEPTUAL

La temática de la **evaluación educativa**, que generalmente ha estado restringida al ámbito del aula de clases, ha tenido en los últimos años una expansión considerable. En el marco general de la reforma de la administración pública, y en el particular de los procesos de reforma educativa orientados a la “descentralización” y al establecimiento de “estándares de rendimiento”, la evaluación educativa viene asumiendo nuevas funciones referidas a la evaluación de macroestructuras (como el sistema educativo nacional, o por niveles o regiones), y la de mesoestructuras (el currículo, proyectos educativos, procesos de formación de docentes, programas e instituciones educativas, etc.) bajo el enfoque de **Accountability o Rendición de Cuentas**. En esta perspectiva se cuenta con una amplia trayectoria en los países anglosajones³, pero es todavía incipiente en la región latinoamericana (Rodríguez, A., 2000:135).

Accountability o Rendición de Cuentas significa fijar la responsabilidad por las propias acciones y los resultados de dichas acciones. Esto implica que debe haber una rendición de cuentas (es decir, una especificación de lo que se ha hecho y los resultados que se han logrado) a algún tipo de autoridad, ya sea un nivel superior en la jerarquía o alguna autoridad externa. En el campo de la educación, esto incluye entregar información a los apoderados, contribuyentes, empleadores y otras partes interesadas por la calidad de la educación. La entrega de información, aunque no define responsabilidades, puede ser útil para identificar problemas y deficiencias y es un paso importante que centra la atención en los aspectos sobre los cuales se pueden tomar acciones para mejorar la educación (McMeekin, 2006).

Siguiendo una tendencia iniciada por la Internacional Association for Educational Achievement (IEA) en los años sesenta (Tiana, 2001), se han multiplicado las iniciativas nacionales, regionales e internacionales que consideran como **principal indicador del desempeño de los sistemas educativos los aprendizajes alcanzados por los estudiantes en pruebas estandarizadas aplicada a muestras representativas de la población estudiantil**. Ejemplos destacados de esta tendencia han sido los estudios de TIMSS (Trends in International Mathematics and Science Study), PIRLS (Progress in International Reading Literacy Study), los estudios del LLECE UNESCO en Latinoamérica y los estudios de PISA (Programme for International Student Assessment), de gran impacto en los contextos sociopolíticos de los países desarrollados en la última década.

³ Sus inicios se ubican en los años setenta en Gran Bretaña. Básicamente se implementó a través de un Currículo Nacional que especifica metas de logro que sirve de base para la evaluación mediante pruebas nacionales de corte estandarizado (McCormick y James, 1996; Terigi, 1997; Rodríguez Ousset, 1997).

Las **pruebas estandarizadas** tienen la particularidad de que contienen los mismos ítems para todos los alumnos que responden a un aplicación temporal dada en un espacio definido (municipio, departamento o país), siendo sus aplicaciones posteriores (usualmente otros años), de dificultad comparable. Usualmente esto se logra mediante un riguroso análisis "de los ítems individuales en pruebas diferentes para asegurar que evalúan las mismas destrezas y tienen una complejidad equivalente (se utilizan ensayos prácticos con muestras de alumnos para verificar que la proporción de respuestas correctas a dos ítems equivalentes sea la misma para alumnos seleccionados al azar). Para estandarizarse de esta manera, las pruebas se basan en respuestas a un conjunto de alternativas, con un número fijo de "distractores" y una sola respuesta "correcta". Las conocidas pruebas de elección múltiple en las que el alumno marca la respuesta correcta entre un conjunto de opciones (a menudo cinco) es el modelo básico de la pruebas estandarizadas." (McMeekin, 2006:32). Las pruebas aplicadas en Honduras se desarrollaron durante dos años a partir del establecimiento de los Estándares Educativos.

Las pruebas estandarizadas pueden estar referidas a norma o a criterio. **Las pruebas referidas a criterio** están orientadas a establecer en qué medida los alumnos evaluados han logrado el dominio de un contenido específico, "de manera que el porcentaje de respuestas correctas indica el grado de cercanía que logra cada alumno (o grupo de alumnos) al "criterio" de dominio de la destreza o conocimiento evaluado. Así, es posible establecer un estándar objetivo o "criterio" respecto del desempeño real del alumno en relación con el desempeño de debería lograr (McMeekin, 2006:35). Este es el tipo de pruebas referidas en el presente estudio.

Las pruebas estandarizadas deben estar **alineadas** con el currículo y los Estándares Educativos, proceso que se desarrolló en Honduras durante los últimos años a partir del DCNB. Como se ha señalado repetidamente en la literatura regional: "Las pruebas utilizadas para evaluar en qué medida los alumnos están logrando el aprendizaje deseado deben estar alineadas tanto con los estándares como con los currículos. No tiene sentido tratar de responsabilizar a las escuelas y alumnos por los resultados específicos a menos que el currículo efectivamente enseñado esté estrechamente relacionado con estos resultados y que las pruebas que evalúan el desempeño están basadas en dicho currículo" (PREAL, 2006:26).

3. MARCO METODOLÓGICO

El diseño de la evaluación siguió el modelo que propone la nueva teoría psicométrica de los test llamada Teoría de Respuesta al Ítem (TRI), que representa un avance significativo en la construcción y análisis de las pruebas estandarizadas. La TRI formula la existencia de una relación entre los valores de la variable que miden los ítems y la probabilidad de dar una respuesta determinada.

Esta teoría supone que el rendimiento académico no es un rasgo observable directamente; sino que es un *rasgo latente*, y establece que la probabilidad de respuesta correcta de dos personas será la misma, si y sólo si, son igualmente hábiles. Plantea que la medida obtenida por una persona no depende de las características de los ítems ni de la puntuación obtenida por otras personas evaluadas. Esta independencia entre el instrumento de medición y las personas es la diferencia esencial entre este enfoque y la Teoría Clásica de Tests (TCT).

La presente evaluación se planificó en base a un proceso de cuatro etapas:

1. Desarrollo de las pruebas: Que consiste en la especificación de las características de la prueba, elaboración y validación de ítems.
2. Producción de pruebas: Impresión, empaque y distribución
3. Administración de pruebas: Diseño de una muestra aleatoria estratificada y representativa, capacitación de aplicadores y aplicación de pruebas.
4. Procesamiento de datos: Digitalización de respuestas de los niños a través de lectores ópticos que escanean las pruebas aplicadas, elaboración de bases de datos, gestión de datos estadísticos para producir reportes.

3.1. Objetivos de la evaluación

La Secretaría de Educación mediante el proceso de “evaluación externa” tiene el **propósito** de valorar el nivel de desempeño alcanzado por los alumnos respecto a los estándares alineados con el DCNB. La evaluación externa es un instrumento objetivo del que se dispone para verificar si los alumnos están aprendiendo lo propuesto en el DCNB, desde pruebas que permiten comparar los resultados de todo un año lectivo a nivel nacional, departamental, etc.

3.1.1. Objetivos específicos

- **Determinar los rendimientos académicos obtenidos en el aprendizaje de español y matemáticas en función de las metas EFA.** Teniendo en cuenta que el Estado de Honduras a través de la Secretaría de Educación asumió el compromiso de transformar la educación en el país; adquiriendo además una serie de compromisos internacionales como las metas del milenio y las metas EFA; considerando que todos estos propósitos se relacionan con el desempeño de los estudiantes, entonces se requiere de la evaluación de aprendizajes a fin de corroborar en qué medida se están haciendo avances en el cumplimiento de las metas.
- **Generar información útil para tomar decisiones que mejoren el desempeño del sistema educativo.** La información producida es de utilidad para la Secretaría de Educación y los distintos actores que intervienen en los procesos educativos, ya que con ella se revisan las acciones formativas a fin de apuntar en la dirección que permita optimizar el rendimiento de la población estudiantil. Los datos generados pueden servir para que a nivel desconcentrado se tomen decisiones locales, también los centros educativos participantes pueden hacer uso de la información para valorar sus alcances y las prácticas que conducen a los resultados obtenidos.

3.2. Aspectos evaluados

El Currículo Nacional Básico acentuó su fase de ejecución en el año 2006 con la distribución de Libros de texto, Estándares Educativos, Programaciones Educativas y Pruebas Formativas; la llegada al aula de estos recursos educativos es acompañada de nuevos enfoques de enseñanza referidos en el DCNB. Esta reforma es un cambio sustantivo e integral, que está alineado y articulado en los distintos niveles del sistema educativo. El presente estudio tiene como punto de partida la evaluación de este nuevo currículo, de manera que se enfoca en establecer en qué medida están siendo alcanzados los conocimientos y habilidades que se proponen en el DCNB.

3.3. Proceso de evaluación

3.3.1. Desarrollo de las pruebas

El punto de partida en la elaboración de las pruebas fueron los Estándares Educativos, que han sido diseñados e implementados por la Secretaría de Educación y que constituyen la concreción de los contenidos que plantea el DCNB, representan el criterio que determina cuáles son los logros que se espera alcancen los estudiantes. Una primera actividad en el diseño de las pruebas fue priorizar y jerarquizar los estándares a ser evaluados.

Teniendo en cuenta que los docentes son los que mejor conocen los contenidos a evaluar, se consideró que los ítems tenían que ser desarrollados por ellos; en este sentido se convocó a un taller a 90 docentes de español y 90 de matemáticas, los cuales fueron seleccionados de los 18 departamentos del país, en base a criterios de idoneidad, tales como experiencia docente y conocimiento del nuevo currículo.

A estos docentes se les brindó una capacitación sobre redacción de reactivos, en la que se les entrenó en actividades para la elaboración de ítems tales como, la búsqueda selectiva de textos y problemas; se les instruyó para que tuvieran mensajes positivos, adecuados al nivel del grado, sin contenido político, religioso o discriminatorio y contextualizado al medio. Se generaron dos tipos de reactivos, los independientes (un enunciado para un reactivo) y los dependientes (un enunciado para varios reactivos). Los reactivos son de selección única, redactados en forma de pregunta, con cuatro opciones; se caracterizan por ser breves, claros, precisos, sin ambigüedad, con una misma extensión, y las opciones incorrectas deberían ser factibles pero claramente incorrectas.

En el taller se elaboraron ítems por grado y asignatura, una vez obtenidos los reactivos se procedió a su revisión, seguidamente, se clasificaron en **ANCLA**, los mejores por su elaboración y representación de los contenidos evaluados; un segundo tipo de ítem son los **BASE**, que son de igual calidad pero que se decide que pueden ser sustituidos en futuras pruebas; ambos conforman los ítems operativos. Los ANCLA son aquellos que tienen la función de permanecer constantes a lo largo de los distintos levantamientos, constituyen alrededor del 15% de la prueba. Los BASE van cambiando de un levantamiento a otro y constituyen un 70% del total de la prueba. Un tercer tipo de ítem son los **NUEVOS**, aquellos que no han sido validados y se aplican por primera vez; el objetivo de su inclusión es probarlos para ir alimentando el banco con ítems validados. Las pruebas que fueron conformadas para la evaluación del 2008 fueron 8 formas diferentes para cada grado.

Antes de que los ítems sean considerados operativos (ANCLA, BASE), deben ser revisados con el propósito de asegurar que estén alineados al estándar y grado correspondiente y que cumplen además varias condiciones: que la respuesta correcta sea única y clara, que los tres distractores sean apropiados y claramente incorrectos, cada distractor representa un error típico de los alumnos del grado en cuestión, las cuatro opciones son de la misma extensión, no hay orientaciones negativas en ninguna de ellas, las respuestas correctas no destacan por ninguna razón, el lenguaje debe ser sencillo, claro, sin ambigüedad, deben estar libres de discriminación en razón de género, étnica o cultural. La revisión de los reactivos fue realizada por equipos técnicos en los que participaron docentes, personal técnico de la Secretaría de Educación, personal de proyectos educativos y profesores universitarios. Requirió de una revisión detallada en la que se detecta si están alineados con los estándares, bloques de contenido y

componentes. Aplicando los criterios antes expuestos, la decisión es aprobarlos, rechazarlos o modificarlos.

- **Banco de ítems.** Es el conjunto de ítems que se tienen para medir el dominio de los contenidos de español y matemáticas, y cuyos parámetros están estimados en una misma escala; permanecen almacenados con el fin de integrar un sistema de evaluación. Ítems pertenecientes al banco están calibrados, es decir, cada uno cuenta con un cálculo de los parámetros de dificultad, discriminación y adivinación. A partir de éstos es posible construir pruebas que proporcionen mediciones más precisas.
- **Mapas de las pruebas.** Es un documento que permite visualizar la ubicación de cada ítem en función del bloque, estándar y componente al que se pertenece. La utilidad de esta herramienta es que posibilita tener una clasificación y organización de los instrumentos que se aplican de tal manera que posteriormente sean analizados con independencia del número de formas que sean administradas.

Ejemplo de esquema del mapa de prueba:

Posición	Forma	Código	Pauta	Bloque	Componente	Estándar
1	1-8	M06M1J01006	4	Números y Operaciones	J. Múltiplos y Divisores	1

3.3.2. Producción de las pruebas

Implica la impresión preliminar de las pruebas para su respectiva revisión y aprobación por parte de los equipos técnicos, después de esto se efectúa una impresión definitiva y se procede al empaque en función del listado de los centros educativos escogidos en la muestra, depositando exclusivamente las pruebas que son requeridas por los alumnos de cada una de las escuelas seleccionadas, las cajas son debidamente rotuladas y selladas, a fin de asegurar su inviolabilidad.

3.3.3. Administración de las pruebas

La aplicación de las pruebas implica la determinación previa de la muestra la cual tiene que ser diseñada, al mismo tiempo debe contarse con un manual para la aplicación de la prueba y efectuarse la capacitación de los aplicadores.

3.3.4. Diseño de la muestra

La evaluación se realizó en base a una muestra de la población de los centros educativos, debido, por un lado, al menor costo de un levantamiento muestral; y también, al tiempo disponible para realizarla. Asimismo, el volumen de información y el grado de control que se tiene en la implementación del proceso muestral, son factores que mejoran la calidad de información recolectada. Considerando que uno de los objetivos es hacer inferencias a nivel nacional y departamental, se requirió diseñar una muestra que fuera representativa, que tuviera el nivel de confianza óptimo y un grado de error aceptable; de manera que responda satisfactoriamente a la necesidad de contar con datos válidos.

Cálculo del tamaño de la muestra

El objetivo de hacer este cálculo, fue establecer un número de centros educativos representativo para hacer inferencias válidas. El punto de partida para la determinación del tamaño es contar con un marco muestral, en este caso el listado de centros educativos que constituyen la población de interés, que fue proporcionado por la UPEG (Unidad de Planificación y Evaluación de la Gestión de la Secretaría de

Educación) en un archivo electrónico en el que aparece la lista de escuelas, nombre, ubicación exacta (departamento, municipio, aldea, caserío, barrio o colonia, calle), tipo de administración (público privado), ubicación (rural urbana), tipo de escuela (uni, bi o multidocente); además del número de alumnos por grado y secciones atendidas. Con esta información se utilizó un programa informático diseñado para calcular la muestra de 854 centros educativos. El detalle de los 854 centros por departamento se presenta en el siguiente cuadro, los cuales aportaron a fines del año lectivo una matrícula global de 101,895 alumnos de primero a sexto grado.

CUADRO N° 4. DETALLE DE CENTROS EDUCATIVOS POR DEPARTAMENTO QUE FUERON EVALUADOS COMO PARTE DE LA MUESTRA NACIONAL

Departamento	N° centros evaluados
Atlántida	39
Choluteca	70
Colón	43
Comayagua	54
Copán	56
Cortés	52
El Paraíso	57
Francisco Morazán	61
Gracias a Dios	25
Intibucá	51
Islas de la Bahía	20
La Paz	73
Lempira	30
Ocotepeque	44
Olancho	43
Santa Bárbara	47
Valle	52
Yoro	40
TOTALES	854

Especificaciones del tipo de muestreo

La muestra utilizada es de tipo estratificado y representativo. Es un tipo de muestreo probabilístico en la medida en que cada elemento de la muestra tiene posibilidad conocida de pertenecer a ésta, lo cual permite conocer la precisión de las características que se estiman. Los estratos son subdivisiones que están presentes en la población (departamento, ubicación geográfica y tipo de centro) y que son tenidos en cuenta en virtud que se quiere ofrecer resultados respecto a los mismos. Con esta estrategia se disminuye la heterogeneidad de los sub-grupos poblacionales y aumenta la exactitud de las características que se miden. Ello es conveniente porque de esta forma se sabrá si las diferencias observadas entre los alumnos pueden ser atribuidas a los estratos establecidos o a factores no determinados.

4. SÍNTESIS DE RESULTADOS Y HALLAZGOS PRINCIPALES

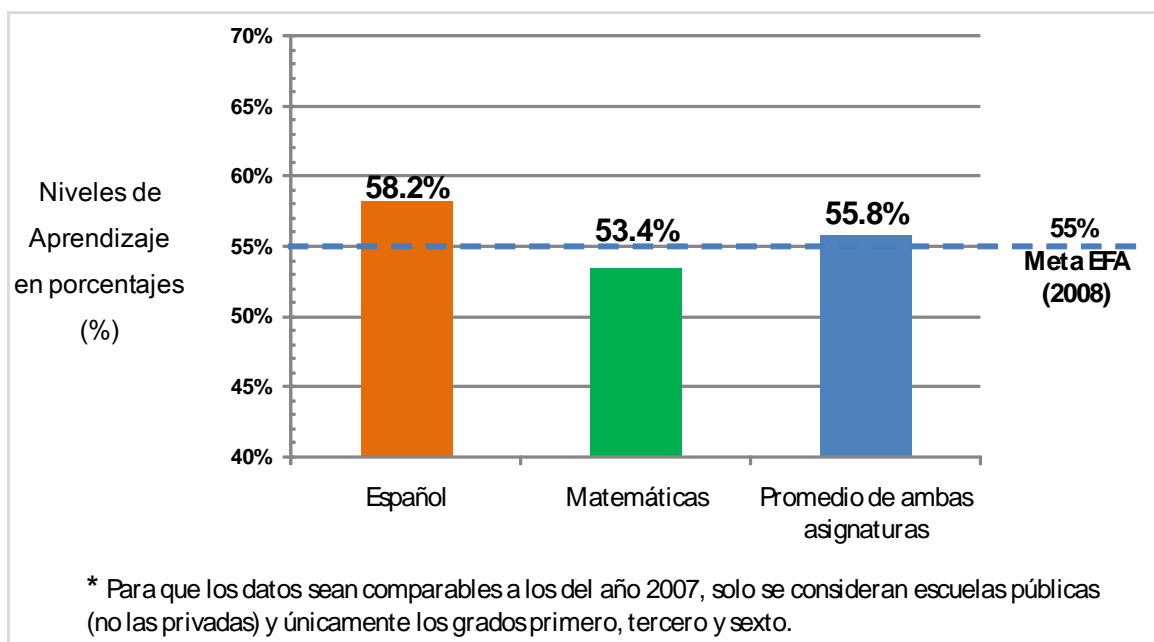
A continuación se presentan los resultados globales más relevantes del presente estudio en relación con el propósito del mismo: Valorar el nivel de desempeño alcanzado por los alumnos respecto a los Estándares Educativos alineados con el DCNB en las asignaturas de Español y Matemáticas de primero a

sexto grado. Se presentan también resultados relacionados con los objetivos “Determinar los Rendimientos Académicos obtenidos en el aprendizaje de Español y Matemáticas en función de las meta EFA” y “Generar información útil para tomar decisiones que mejoren el desempeño del sistema educativo” a través del análisis comparativo de categorías tales como resultados por grado y asignatura, departamentos, urbano-rural, uni-bi-multidocente, pública-privada y masculino-femenino. El análisis detallado de logro por Estándar Educativo por asignatura y grado no se presenta en este resumen.

Los resultados promedio en las escuelas públicas, en Español y Matemáticas, en primero, tercero y sexto grado⁴, indican que, por primera vez desde que se establecieron, se ha alcanzado la meta EFA de Rendimiento Académico para el año correspondiente (2008: 55%) en la asignatura de Español, aunque no en Matemáticas.

Los valores globales de la muestra nacional de alumnos de centros educativos públicos para primero, tercero y sexto grado en la asignatura de Español, superan por primera vez desde que se establecieron, la meta EFA en Rendimiento Académico para un año específico: 58.2% respecto a 55% que era la correspondiente al año lectivo 2008. Sin embargo, los resultados promedio para los mismos alumnos en la asignatura de Matemáticas se quedan debajo de la meta: 53.4% respecto al 55% establecido previamente como meta anual de EFA. Es destacable que, el promedio de las dos asignaturas supera por primera vez la meta EFA anual para los aprendizajes de los alumnos: 55.8%, ocho décimas arriba de la meta. El siguiente gráfico ilustra la relación entre los resultados obtenidos para el año lectivo 2008 y la meta EFA correspondiente.

GRÁFICO 2. RESULTADOS DE APRENDIZAJE EN ESCUELAS PÚBLICAS, PARA PRIMERO, TERCERO Y SEXTO GRADO, EN ESPAÑOL Y MATEMÁTICAS *, COMPARADA CON META EFA PARA EL AÑO 2008



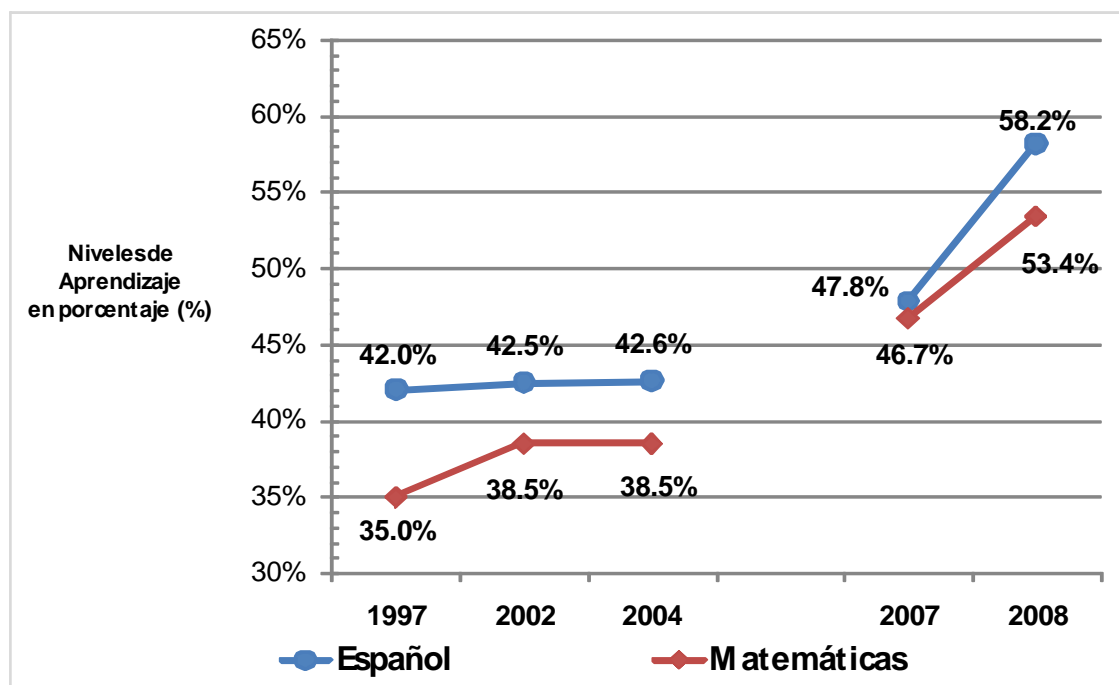
⁴ Para que los datos sean comparables a los del año recién pasado 2007, los resultados del 2008 para evaluar el logro de las metas EFA se limitan a primero, tercero y sexto grado en escuelas públicas, pese a que en este último año lectivo 2008 la muestra incluyó los seis grados de Primer y Segundo Ciclo de Educación Básica y también centros públicos y privados.

FUENTE: Elaboración propia en base a MIDEH/DIGECE. 2009

Pese a que técnicamente no son comparables, una mirada diacrónica a los resultados de evaluación externa de la última década (1997-2008), permite apreciar una mejoría importante en los resultados de aprendizaje en Español y Matemáticas.

Aún cuando las evaluaciones del 2007 y 2008 están basadas en los Estándares Educativos alineados al DCNB, mientras que las aplicadas en los años 1997, 2002 y 2004 se basaban en los Rendimientos Básicos e Indicadores de Evaluación (R.B.), es inevitable una "mirada" comparativa en los niveles de logro alcanzados por los alumnos de la educación básica hondureña, en este caso de las escuelas públicas ya que solamente en este tipo de centros se realizó la evaluación en el año 2007. Y en esta perspectiva, es apreciable una tendencia ascendente en ambas asignaturas. En Español se ha pasado de valores de 42%, 42.5% y 42.6% para los años 1997, 2002 y 2004 respectivamente, a resultados de 47.8% y 58.2% en los años 2007 y 2008. Similar tendencia muestran los resultados de Matemáticas, pasando de valores de 35%, 38.5% y 38.5% para los años 1997, 2002 y 2004, a resultados de 46.7% y 53.4% para 2007 y 2008 respectivamente. La mejoría en los últimos años es significativa y es coincidente con la distribución y uso de todo un conjunto de materiales educativos de apoyo al trabajo de aula: Libros de texto para Español y Matemáticas, Guías didácticas para docentes, Cuadernos de trabajo para alumnos, Estándares Educativos, Programaciones Educativas y Pruebas Formativas Mensuales. La correlación positiva entre la disposición y el uso de estos materiales en el aula, con los resultados de aprendizaje en las pruebas de Fin de Grado del 2007, ha sido verificada en análisis previos (SE, 2008).

GRÁFICO 3. RENDIMIENTO ACADÉMICO PROMEDIO EN ESPAÑOL Y MATEMÁTICAS EN RESULTADOS DE EVALUACIÓN EXTERNA POR AÑO (*)



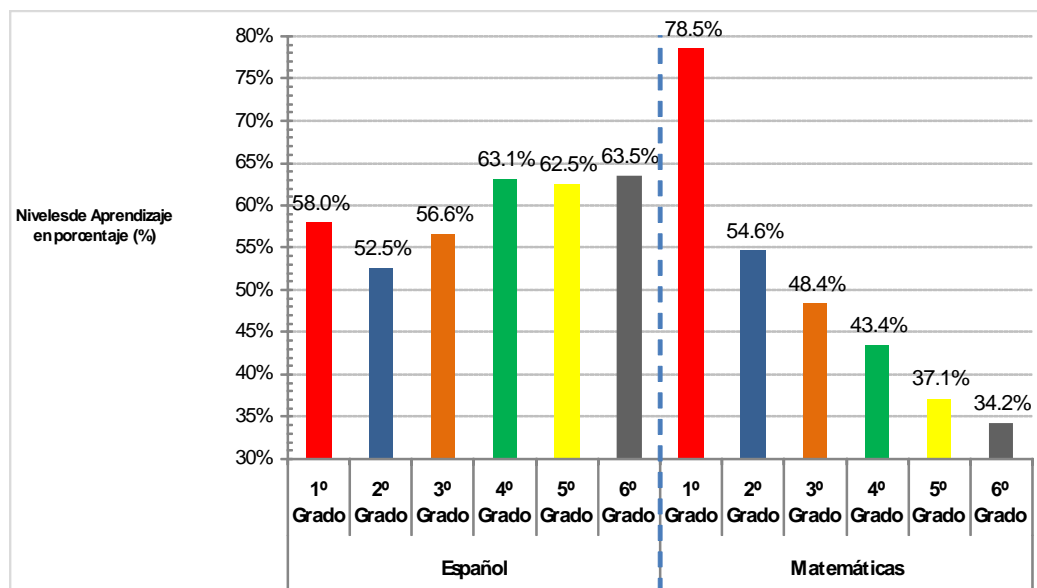
(*) Técnicamente los resultados de 1997/ 2002/ 2004 no son comparables a los correspondientes a los años 2007 y 2008 debido al cambio en el currículo y grados evaluados, por ello los gráficos están cortados entre el 2004 y el 2007. Los resultados del período 1997-2004 están basados en los Rendimientos Básicos e Indicadores de Evaluación (R.B.) para 3ero y 6to grado, mientras que los del 2007 y 2008 se basan en Estándares Educativos alineados con el nuevo Currículo de

Educación Básica y coinciden con la reciente distribución y uso del conjunto de materiales educativos de apoyo a la labor docente, y en el caso de estos años para 1ero, 3ero y 6to grado.

Los resultados promedio en la asignatura de Español son relativamente homogéneos entre los alumnos de los seis (6) grados evaluados en el último año (2008) (mínimo: Segundo grado con 52.5% y máximo 63.5% en Sexto grado), mientras que los resultados correspondientes en Matemáticas presentan una mayor heterogeneidad y una acentuada tendencia descendente desde primer grado hasta el sexto (mínimo: Sexto grado con 34.2% y máximo Primero con 78.5%).

Es destacable que los resultados en la asignatura de Español, pese a las dificultades implicadas por la nueva perspectiva metodológica, el enfoque comunicativo, muestran una tendencia relativamente homogénea en sus resultados desde primero a sexto grado. Por el contrario, los resultados de Matemáticas muestran una varianza muy grande entre los promedios de los diferentes grados y además, una marcada tendencia descendente desde valores de 78.5% y 54.6% en primero y segundo grado, hasta promedios de 37.1% y 34.2% en quinto y sexto grado respectivamente. Estudios previos han hecho referencia a las dificultades que enfrentan los docentes respecto al nuevo enfoque metodológico propuesto por el DCNB para la enseñanza de la asignatura, sus deficiencias de conocimientos en varios de los "nuevos" temas que incluye el DCNB para Matemáticas, y su insatisfacción con las capacitaciones recibidas.

GRÁFICO 4. RESULTADOS COMPARATIVOS DE ESPAÑOL Y MATEMÁTICAS, SEGÚN GRADO. HONDURAS. 2008



FUENTE: Elaboración propia en base a MIDEH/DIGECE. 2009

Mientras en la asignatura de Español casi seis de cada diez alumnos cumplen la Meta EFA para el año 2008, en Matemáticas menos de cuatro de cada diez alumnos alcanzan dicha Meta de Aprendizaje.

Los resultados en Español no sólo son más elevados que los de Matemáticas en el promedio de grado, sino que los superan ampliamente en relación con la proporción de alumnos de la muestra que alcanzaron la Meta EFA de Aprendizaje para el año 2008 (55% en cada asignatura). En promedio, el 57% de los alumnos alcanzaron o superaron dicha meta en Español, con el 66.9 de sexto grado, en esa

condición. Únicamente en segundo grado, menos de la mitad de los estudiantes evaluados alcanzó o superó la meta EFA.

CUADRO 5. RESULTADOS DE EVALUACIÓN EN ESPAÑOL POR GRADO Y PORCENTAJE DE LOS ALUMNOS QUE ALCANZAN O SUPERAN LA META EFA DE APRENDIZAJE. HONDURAS.2008

Grado	Promedio de grado	Proporción de alumnos que alcanzan la Meta EFA (55%)
Primero	58.0%	51.1%
Segundo	52.5%	43.2%
Tercero	56.6%	53.0%
Cuarto	63.1%	64.9%
Quinto	62.5%	63.6%
Sexto	63.5%	66.9%

FUENTE: Elaboración propia en base a MIDEH/DIGECE. 2009

Los resultados de Matemáticas en cambio son significativamente inferiores a los de Español, no sólo en los promedios por grado sino en la proporción de alumnos que alcanzan la Meta EFA de Aprendizaje para el año 2008 (55%). Solamente en el primer grado ocurre que la mayoría de los alumnos alcanzan dicha meta, en el resto es una minoría la que logra esos niveles de aprendizaje, con promedios menores a uno de cada diez alumnos en quinto y sexto grado.

CUADRO 6. RESULTADOS DE EVALUACIÓN EN MATEMÁTICAS POR GRADO Y PORCENTAJE DE LOS ALUMNOS QUE ALCANZAN O SUPERAN LA META EFA DE APRENDIZAJE PARA EL AÑO 2008

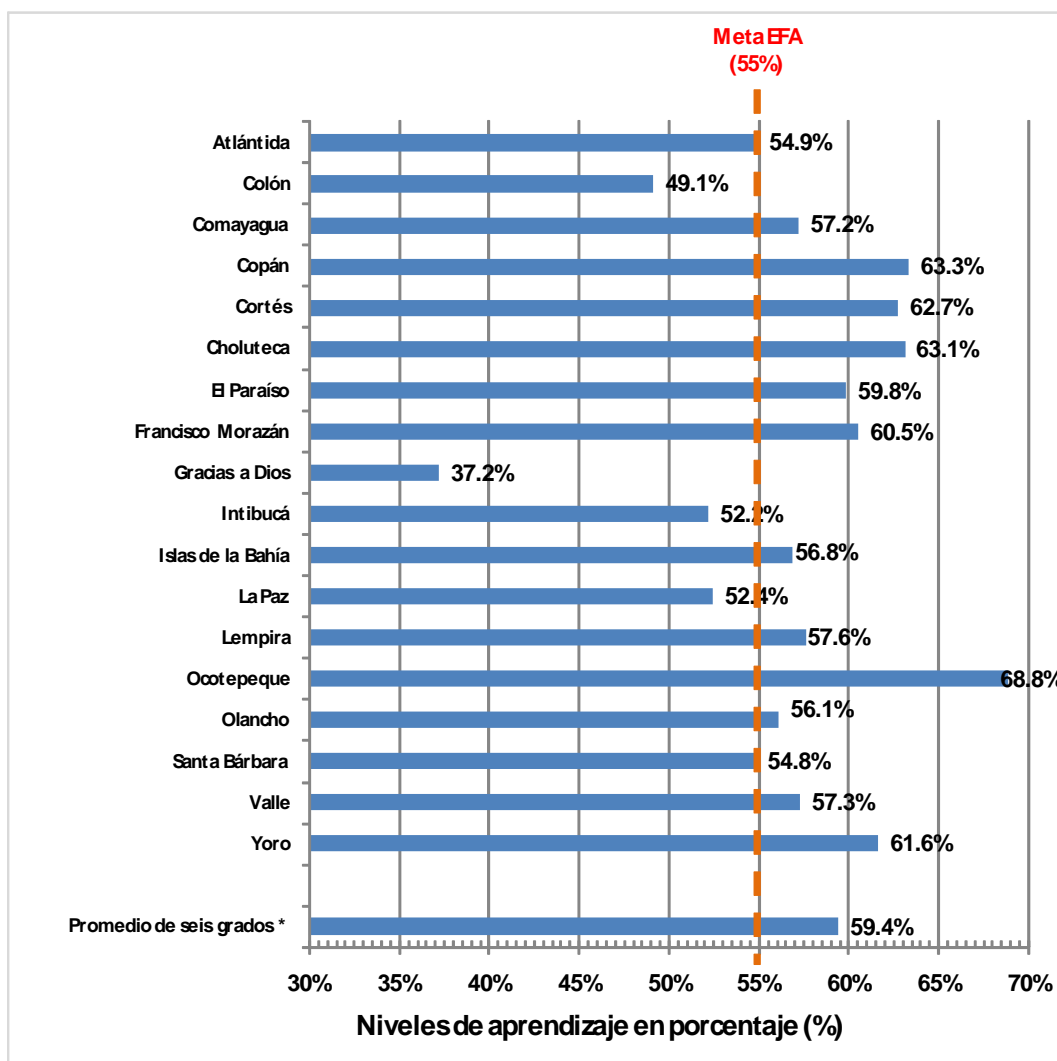
Grado	Promedio de grado	Proporción de alumnos que alcanzan Meta EFA (55%)
Primero	78.5%	88.1%
Segundo	54.6%	49.7%
Tercero	48.4%	34.4%
Cuarto	43.45	23.3%
Quinto	37.15	11.75
Sexto	34.2%	7.7%

FUENTE: Elaboración propia en base a MIDEH/DIGECE. 2009

Existen diferencias significativas entre los resultados de aprendizaje de los alumnos de diferentes departamentos, tanto en Español como en Matemáticas. Ocotepeque, Copán y Choluteca destacan con los promedios más altos en ambas asignaturas, mientras que Gracias a Dios, Colón, Intibucá y Atlántida presentan los resultados más deficitarios.

El análisis de los resultados a nivel departamental ofrece información valiosa para los directivos de esas instancias regionales. Los promedios alcanzados por los alumnos (la muestra es estadísticamente representativa a nivel departamental) indican claras diferencias en los niveles de aprendizaje alcanzados entre los diferentes departamentos. Es significativo que, ni Francisco Morazán ni Cortés aparezcan entre los tres primeros lugares ni en Matemáticas ni en Español. Por el contrario, Ocotepeque presenta los resultados más altos en ambas asignaturas (en las evaluaciones desarrolladas en el año 2004, este departamento también alcanzó el primer lugar), Copán aparece en segundo lugar también en Matemáticas y Español, lo que constituye un cambio sustancial respecto a los resultados del 2004, en el cual aparecía en el grupo de menor aprendizaje para tercer grado. En tercer lugar en ambas asignaturas está Choluteca. El detalle por departamento se presenta en el siguiente gráfico, en el que puede identificarse cuáles departamentos están sobre la Meta EFA de Aprendizaje para el año 2008 (55%), cuáles en la meta, y cuáles bajo la meta.

GRÁFICO 5. RESULTADOS PROMEDIO DE LOS ALUMNOS DE LOS SEIS GRADOS PARA ESPAÑOL, POR DEPARTAMENTO. HONDURAS. 2008



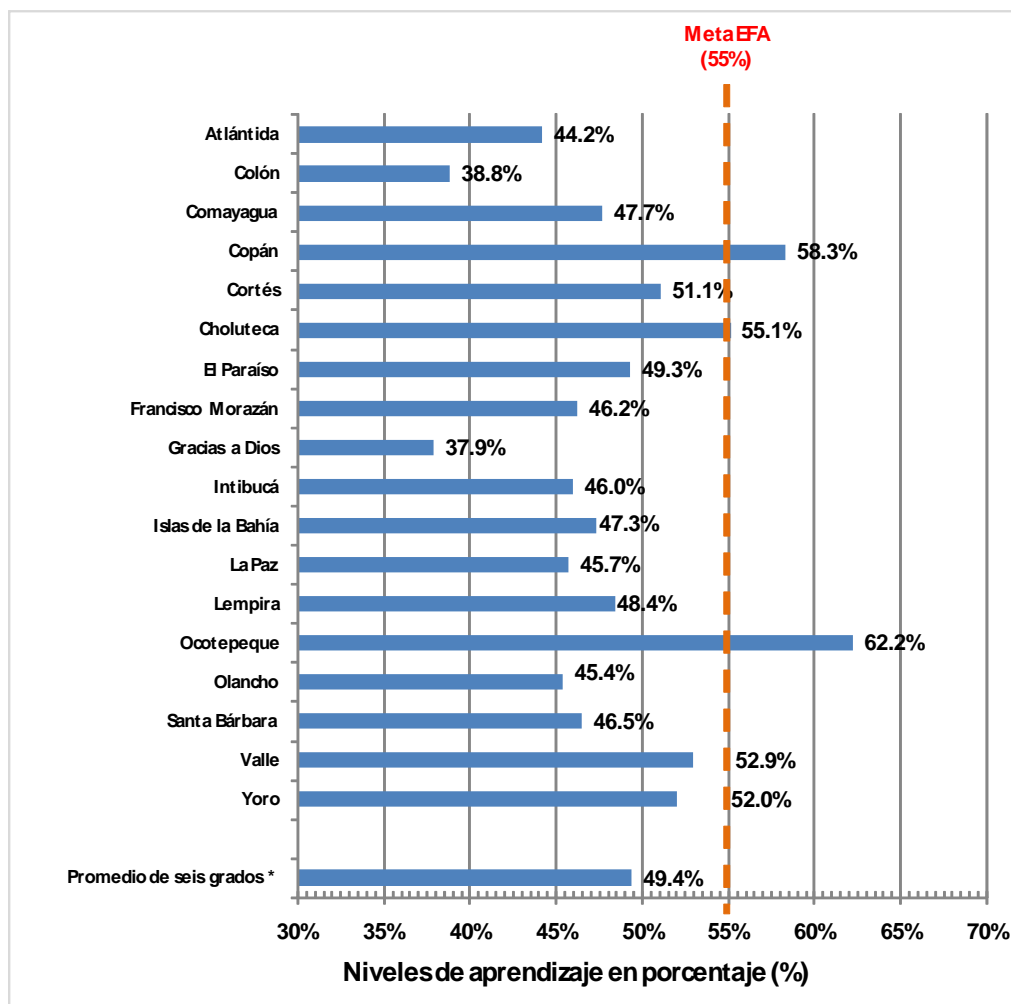
* Este promedio nacional de los seis grados difiere del presentado en el primer hallazgo, porque éste último se refiere sólo a primero, tercero y sexto grado de escuelas públicas.

FUENTE: Elaboración propia en base a MIDEH/DIGECE. 2009

Los resultados alcanzados por los alumnos en Matemáticas fueron notoriamente inferiores, de manera que únicamente en los departamentos de Ocotepeque, Copán y Choluteca, los promedios superan la Meta EFA de Aprendizaje para el año 2008. Resultados que se vuelven más notables si se considera que Ocotepeque y Colón se ubican entre los departamentos de menor desarrollo socioeconómico, cuyos valores de Índice de Desarrollo Humano (IDH) están entre los cinco más bajos del país. Por el contrario Francisco Morazán, Islas de la Bahía y Cortés, los tres departamentos de mayor IDH del país, presentan

resultados de aprendizaje bastante debajo de la Meta EFA. El detalle de los promedios departamentales en Matemáticas se presenta a continuación.

GRÁFICO 6. RESULTADOS PROMEDIO DE LOS ALUMNOS DE LOS SEIS GRADOS PARA MATEMÁTICAS, POR DEPARTAMENTO. HONDURAS. 2008



* Este promedio nacional de los seis grados difiere del presentado en el primer hallazgo, porque éste último se refiere sólo a primero, tercero y sexto grado de escuelas públicas.

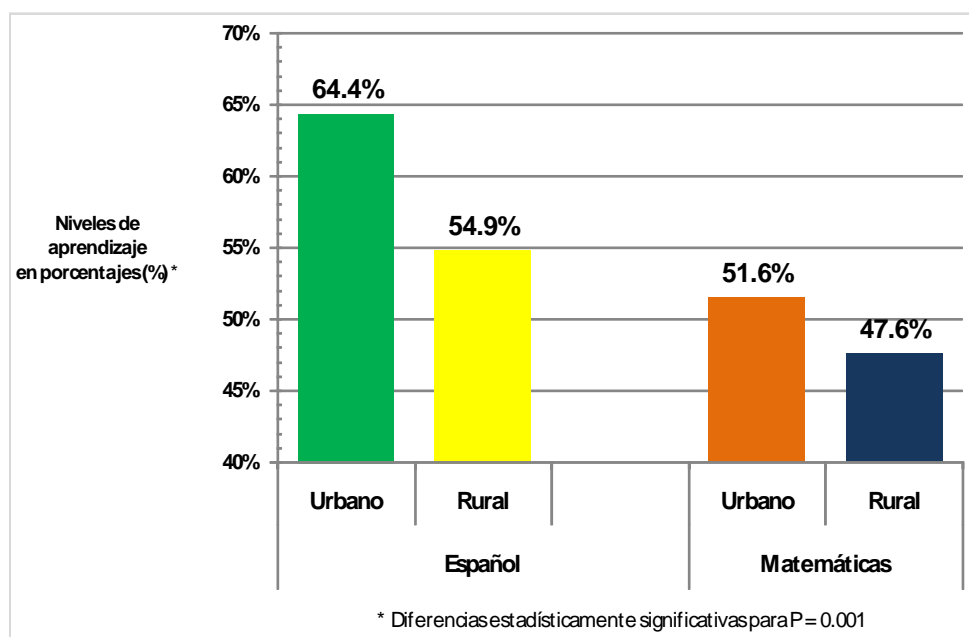
FUENTE: Elaboración propia en base a MIDEH/DIGECE. 2009

Los resultados promedio de aprendizaje alcanzados por los alumnos cuyos centros educativos se ubican en el área urbana son superiores (diferencia estadísticamente significativa), a los resultados de los alumnos de área rurales, tanto en Matemáticas como en Español, en los seis grados evaluados.

Los niveles de aprendizaje evidenciados por los alumnos cuyos centros educativos se ubican en áreas urbanas son significativamente superiores a los correspondientes de las áreas rurales. Las diferencias son más marcadas en Español (en segundo, cuarto y quinto grado la diferencia supera los diez puntos porcentuales) y bastante menores en Matemáticas. Estos resultados son coherentes con los mostrados en las cuatro experiencias de evaluación externa citadas anteriormente (1997, 2002, 2004 y 2007), lo que

puede interpretarse como la persistencia de una marcada inequidad en las oportunidades equitativas entre ambos grupos.

GRÁFICO 7. NIVELES DE APRENDIZAJE PROMEDIO PARA LOS SEIS GRADOS, SEGÚN ASIGNATURA Y ÁREA GEOGRÁFICA DE UBICACIÓN DEL CENTRO EDUCATIVO. HONDURAS. 2008

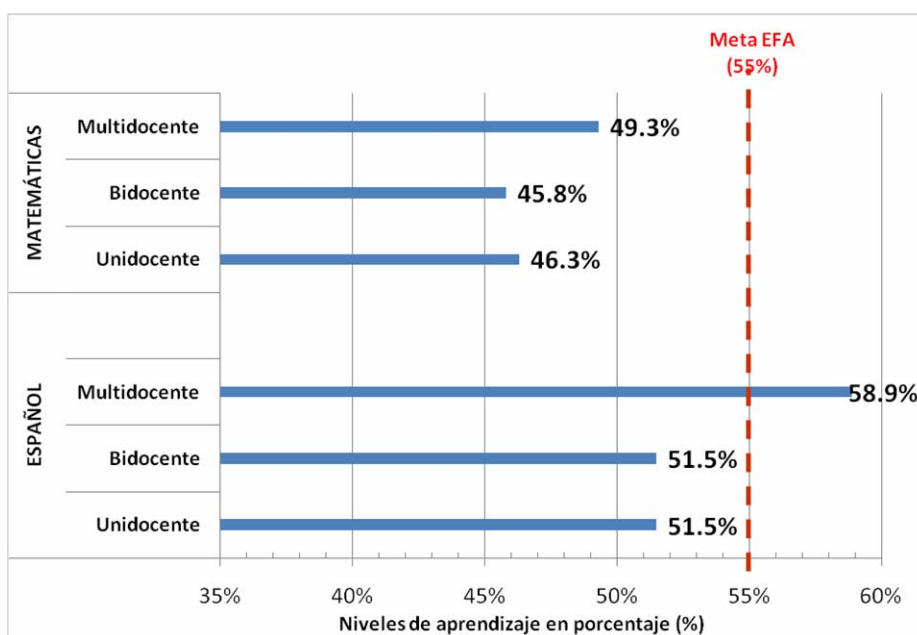


FUENTE: Elaboración propia en base a MIDEH/DIGECE. 2009

Los resultados promedio de aprendizaje de los alumnos de escuelas “multidocentes” (tres o más docentes en el centro), son superiores a los de los alumnos de escuelas “uni” y “bidocentes”, tanto en Matemáticas como en Español, en los seis grados evaluados.

Dado que la existencia de escuelas uni y bidocentes es un fenómeno típicamente rural, la diferencia entre los aprendizajes de las escuelas “Multidocentes” (con tres o más docentes en el centro) y aquellas, puede considerarse otra modalidad que asume la inequidad entre las oportunidades educativas de los alumnos del área urbana respecto a los del área rural. Un detalle de especial significado es el hecho de que, en los promedios por categorías, no aparecen diferencias estadísticamente significativas entre los alumnos de las escuelas unidocentes y los de centros bidocentes, ni en Matemáticas ni en Español.

GRÁFICO 8. NIVELES DE APRENDIZAJE PROMEDIO PARA LOS SEIS GRADOS, SEGÚN ASIGNATURA Y TIPO DE ESCUELA. HONDURAS. 2008

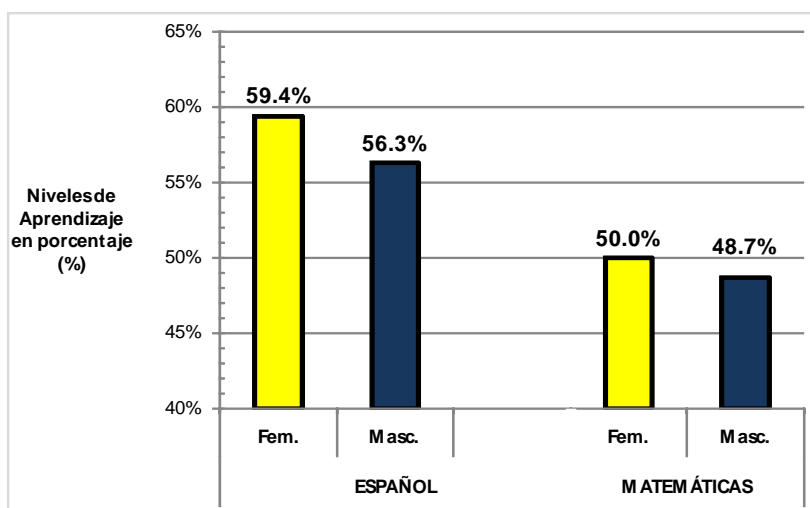


FUENTE: Elaboración propia en base a MIDEH/DIGECE. 2009

Los resultados promedio por género muestran niveles de aprendizaje superiores para las alumnas tanto en Matemáticas como en Español y en los seis grados evaluados.

Es destacable que los resultados promedio de todos los grados en las asignaturas evaluadas muestran diferencias a favor de las alumnas. Dichas diferencias son estadísticamente significativas para el caso de Español (para $P = 0.01$). Los resultados parecen indicar que no se presentan problemas de discriminación por género al interior de los centros educativos, al menos en términos de oportunidades de aprendizaje.

GRÁFICO 9. NIVELES DE APRENDIZAJE PROMEDIO DE LOS SEIS GRADOS, SEGÚN ASIGNATURA Y GÉNERO DE ALUMNADO. HONDURAS. 2008

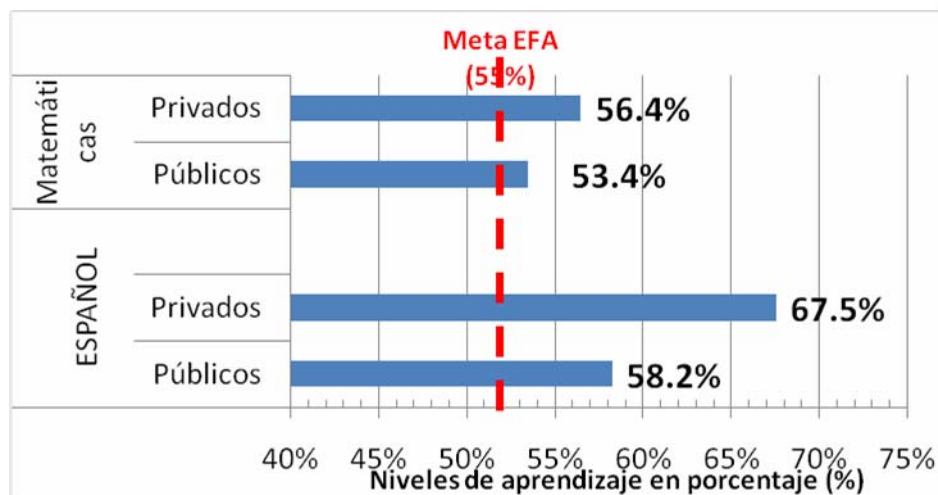


FUENTE: Elaboración propia en base a MIDEH/DIGECE. 2009

Los resultados promedio de aprendizaje de los alumnos de centros educativos de administración privada son superiores a los de los centros públicos, tanto en Español como en Matemáticas.

Los niveles de aprendizaje mostrados por los alumnos de centros privados superan en ambas asignaturas, Matemáticas y Español, a los alumnos de centros públicos. Las diferencias son más marcadas en Español con valores de hasta diez puntos porcentuales en tercero y sexto grado, mientras que en Matemáticas son significativamente menores, de solo uno o dos puntos porcentuales en quinto y sexto grado.

GRÁFICO 10. NIVELES DE APRENDIZAJE PROMEDIO DE LOS SEIS GRADOS, SEGÚN TIPO DE ADMINISTRACIÓN DEL CENTRO EDUCATIVO



FUENTE: Elaboración propia en base a MIDEH/DIGECE. 2009

Los diez centros educativos cuyos alumnos alcanzan, en promedio de los seis grados y las dos asignaturas, los mejores resultados de la muestra evaluada, se ubican en pequeñas comunidades, la gran mayoría rurales.

Es significativo que ninguna de las diez escuelas con los mayores niveles de aprendizaje promedio en Español y Matemáticas y en los seis grados, se ubica en una ciudad importante del país, ni en los

departamentos de mayor desarrollo socioeconómico. La mayoría de estos diez centros son en su mayoría del área rural con solamente uno o dos docentes que atienden 6 ó 3 grados, según sea la escuela unidocente o bidocente, respectivamente. Cuatro de los diez se ubican en Ocotepeque, dos de ellos en el mismo municipio de Santa Fe, dos en Copán, dos en Choluteca, una en Yoro y una en Olancho. El departamento de Ocotepeque ya ha ocupado el primer lugar a nivel nacional en dos evaluaciones externas previas, por lo que merece un estudio especial, en particular sus distritos con “Escuelas Eficaces”.

CUADRO 7. LISTADO DE LOS DIEZ CENTROS EDUCATIVOS CON MEJORES RESULTADOS DE APRENDIZAJE PROMEDIO EN LOS SEIS GRADOS EN LAS DOS ASIGNATURAS DE LA MUESTRA EVALUADA NIVEL NACIONAL HONDURAS. 2008

Departamento	Código	Nombre del centro	Municipio	Lugar	Tipo	Admon.	Ubicación	Promedio
Ocotepeque	141400014	Elisa Cisneros de Carías	Santa Fe	San Cayetano	Bidocente	Pública	Rural	82.55
Ocotepeque	141400014	Manuel María Fernandez	Santa Fe	El Sunnete	Bidocente	Pública	Rural	81.91
Copán	41200030	José Antonio Trigueros	La Unión	La Laguna	Unidocente	Pública	Rural	81.19
Yoro	180300081	PROHECO Rafael Martínez	El Negrito	Nueva Arcadia	Bidocente	Pública	Rural	81.06
Choluteca	60100098	Adriana Espinal de Argeñal	Choluteca	Hato Nuevo N.2	Multidocente	Pública	Rural	81.03
Ocotepeque	140700013	PROHECO José Antonio Arrollas	La Labor	El Azufrado	Bidocente	Pública	Rural	80.21
Olancho	151000037	Rafael Leonardo Callejas	Guata	San Diego	Bidocente	Pública	Rural	80.16
Choluteca	61300005	Juan Lindo	San Isidro	Los Amates	Multidocente	Pública	Rural	78.97
Ocotepeque	141600030	Renovación Cristiana	Sinuapa	Sinuapa	Multidocente	Privada	Urbana	78.17
Copán	41200005	El Progreso	La Unión	Los Arroyos	Multidocente	Pública	Rural	77.83

Fuente: Elaboración propia en base a MIDEH/DIGECE. 2009

5. IMPLICACIONES DE LOS RESULTADOS

Los resultados de evaluación externa son siempre una rica fuente de información para los actores del sistema educativo y para los tomadores de decisiones. En el contexto de una educación hondureña en estado crítico, con sus indicadores educativos ubicados entre los más bajos de la región, debe recordarse en primer lugar que la Evaluación estandarizada aporta información fundamental e indispensable sobre cómo está operando el sistema, pero no es un indicador completo del mismo; y en segundo lugar, que **la evaluación es una condición necesaria, pero no suficiente para mejorar la educación**, aporta información, pero las acciones que se tomen a partir de ella es lo que puede incidir en la mejora de la educación. En este sentido, el “Informe Nacional de Evaluaciones de los Aprendizajes (2008)” aporta valiosa información para diferentes niveles de acción.

A un nivel **Macro**, la Secretaría de Educación recibe información sobre los logros de aprendizaje que los alumnos están alcanzando en relación con el currículo prescrito, por asignatura, bloques de contenido, estratos por región geográfica, tipo de escuela, etc. Lo cual constituye un insumo de primer nivel para el diseño o reorientación de política educativa, programas o proyectos, sean de dimensión nacional o

regional. La evaluación externa debería convertirse en un insumo habitual de los tomadores de decisiones a nivel central. Por ejemplo, la Secretaría de Educación puede elaborar estrategias puntuales para atender deficiencias de aprendizaje observadas en los resultados, ya que éstos ofrecen un importante insumo sobre cuáles son los principales bloques de contenido y estándares en los que están fallando los estudiantes en Matemáticas y Español, y a partir de lo cual se pueden proponer acciones concretas para reforzar el aprendizaje de los estudiantes. En concreto, de acuerdo a los resultados aquí presentados, la capacitación debe enfocarse en mejorar las competencias docentes para atender la enseñanza de Matemáticas en el segundo ciclo, porque los resultados son significativamente más reducidos que en primer ciclo. En cambio, los esfuerzos de capacitación en Español deben ser enfatizados en el de primer ciclo ya que revelan resultados menos favorables.

A nivel **Meso**, las Direcciones Departamentales y Distritales disponen de insumos relevantes y específicos sobre la situación de sus centros educativos, comparable tanto a nivel interno como extra regional. Estas instancias pueden elaborar un plan de seguimiento, supervisión y **acompañamiento pedagógico**, teniendo en cuenta los hallazgos de los respectivos departamentos, con el propósito de continuar mejorando para alcanzar las metas propuestas. Pueden apoyarse en los logros y las dificultades educativas presentes en el conjunto del sistema, para estudiar en profundidad las debilidades propias que influyen en el aprendizaje. Desde esta perspectiva pueden mejorar su labor de orientación a los docentes. En particular los que realizan labores de supervisión y/o acompañamiento pedagógico pueden beneficiarse al contar con un instrumento de apoyo para su labor. El análisis de cuáles bloques de contenido y estándares educativos son en los que más fallan los estudiantes del departamento/ distrito permite el diseño de un plan de capacitación específico, relevante y pertinente para los docentes implicados.

A un nivel **Micro**, la evaluación externa presenta valiosa información específica de los centros educativos evaluados. El director y los docentes del centro educativo pueden analizar los reportes de resultados por escuela, a fin de que los mismos sirvan para revisar la planificación, ya que se dispone en los informes por escuela de datos acerca del nivel alcanzado en Matemáticas y Español, pero también por bloques de contenido, lo que posibilita comprender específicamente dónde están las limitaciones para alcanzar las metas planteadas. En su caso, la evaluación constituye una mirada externa que con objetividad valora lo que están logrando y lo que no están logrando los propios estudiantes, cómo están aprendiendo y qué dificultades están teniendo; además pueden compararse con otros centros educativos de su entorno y a nivel nacional con la intención de extraer lecciones acerca de “como lo están haciendo los demás”. El análisis de los resultados amerita reuniones colectivas de los docentes del centro educativo a fin de que puedan reflexionar acerca de los hallazgos obtenidos para el centro y poder establecer compromisos colectivos para lograr una mejora en la calidad de los resultados. La información también es relevante para los padres de familia, pues disponen de una fuente externa que les indique cuánto están aprendiendo sus hijos en relación con lo propuesto en el currículo, además de enterarse sobre cómo está la situación del centro educativo de su comunidad respecto a otros de la zona, región, departamento o incluso a nivel nacional.

Finalmente, debe recordarse que **“la evaluación estandarizada externa sólo tendrá efectos positivos sobre la educación, si es concebida, percibida y empleada como un mecanismo de responsabilización pública de todos los actores vinculados al quehacer educativo”, para orientar la toma de decisiones en cada nivel de acción** (PREAL, 2008:9).

REFERENCIAS BIBLIOGRÁFICAS

- Bertoni, A., Poggi, M. y Teobaldo, M. (1995). *Evaluación. Nuevos significados para una práctica compleja*. Buenos Aires: Kapelusz.
- Carnoy, M, Elmore, R. y Siskin, L. (2003). *The new accountability*. New York: Routledge Falmer.
- CLEB (2001). *Procesos de Reforma Educativa*. Santiago de Chile: CLEB.
- Edel, R. (2003). El rendimiento académico: concepto, investigación y desarrollo. *Revista Electrónica Iberoamericana sobre Calidad, Eficacia y Cambio en Educación*, 1(2). <http://www.ice.deusto.es/rinace/reice/vol1n2/Edel.pdf>. Consultado el 15 enero 2009.
- Ferrer, G. (2006). *Sistemas de evaluación de aprendizajes en América Latina. Balance y desafíos*. Santiago de Chile: PREAL.
- Garrido, J.L. (1995). *Problemas Mundiales de la Educación*. Madrid: Dykinson.
- Hambleton, R.K. Swaminathan, H. y Rogers, H. (1991). *Fundamentals of Item Theory*. Newbury Park: Sage.
- Hambleton, R. K. (1997). Perspectivas futuras y aplicaciones. En: J. Muñiz, *Introducción a la Teoría de Respuesta a los Ítems*. Madrid: Ediciones Psicología Pirámide.
- LLECE - Laboratorio Latinoamericano de Evaluación de la Calidad de la Educación (2001). *Primer estudio internacional comparativo sobre Lenguaje, Matemática y factores asociados, para alumnos del tercer y cuarto grado de la educación básica*. Santiago de Chile: UNESCO.
- LLECE - Laboratorio Latinoamericano de Evaluación de la Calidad de la Educación (2008). *Segundo estudio regional comparativo y explicativo. Los aprendizajes de los estudiantes de América Latina y el Caribe*. Santiago de Chile: UNESCO.
- Lord, Frederic M. (1980). *Applications of Item Response Theory to Practical Testing Problems*. New Jersey: Lawrence Erlbaum Associates.
- Martínez Arias, R. (1997). *Psicometría. Teoría de los Tests Psicológicos y Educativos*. Madrid: Ediciones Síntesis.
- Masters, G. y Keeves, J. (1999) *Advances in Measurement in Educational Research and Assesment*. Oxford: Pergamon.
- McCormick, R. y James, M. (1996). *Evaluación del curriculum en los centros escolares*. Madrid: Morata.
- McMeekin, R. (2006). *Hacia una comprensión de la Accountability Educativa y como puede aplicarse en América Latina*. Santiago de Chile: Merino.
- Murillo, F.J. (2005). *Estudios sobre Eficacia Escolar en Iberoamérica. 15 buenas investigaciones*. Bogotá: Convenio Andrés Bello.
- Murillo, F.J. (2007). School Effectiveness Research in Latin America. En T. Townsend (Ed.), *International Handbook of School Effectiveness and Improvement* (pp. 75–92). New York: Springer.
- Nirenberg, Brawerman y Ruíz. (2000). *Evaluar para la transformación*. Buenos Aires: Paidós.
- PREAL (2008). *Las evaluaciones educativas que América Latina Necesita*. Santiago de Chile: PREAL.
- PREAL (2006). *Accountability educacional: Posibilidades y desafíos para América Latina a partir de la*

experiencia internacional. Santiago de Chile: San Marino.

PREAL – FEREMA (2006). *Informe de progreso educativo: Honduras.* Honduras: PREAL.

PREAL (2003). *Entre el avance y el retroceso.* Santiago de Chile: PREAL.

PREAL (2001). *Quedándonos atrás. Un informe del progreso educativo en América Latina.* Santiago de Chile: PREAL.

Reimers, F. (1990). Deuda externa, ajuste estructural y educación en América Latina. Tiempo de crisis y reformas. *Revista Latinoamericana de Estudios Educativos.*

Rodríguez, A. (2000). La evaluación de dispositivos educativos. *Revista Pedagógica Universitaria.*

Schiefelbein, E. y Schiefelbein, P. (2008). Evolución de los Procesos de Evaluación del Sistema Educativo 1950-2008. *Revista Iberoamericana de Evaluación Educativa*, 1 (1), pp. 45-50. <http://www.rinace.net/riee/numeros/vol1-num1/art3.pdf>. Consultado el 15 de enero 2009.

Schmelkes, S. (1996). La evaluación de los Centros Escolares. *Taller sobre evaluación de Docentes y Centros Educativos.* Cancún Quintana Roo, México.

LA EVALUACIÓN DE LA EDUCACIÓN CÍVICA Y ÉTICA EN MÉXICO. UN RECORRIDO POR 3º DE PRIMARIA Y 3º DE SECUNDARIA

Oscar Barrera

Revista Iberoamericana de Evaluación Educativa 2009 - Volumen 2, Número 1

<http://www.rinace.net/riee/numeros/vol2-num1/art6.pdf>

Fecha de recepción: 13 de marzo de 2009

Fecha de comunicación de dictamen: 14 de abril de 2009

Fecha de aceptación: 14 de abril de 2009

1. LA EVALUACIÓN DE LA EDUCACIÓN CÍVICA Y LA FORMACIÓN CÍVICA Y ÉTICA EN MÉXICO

La formación de ciudadanos ha sido una tarea fundamental desde la época helénica hasta nuestros días. La integración de niños y jóvenes que compartan la educación de las costumbres del colectivo para un desarrollo equilibrado en la sociedad brindan la posibilidad de una mayor cohesión social. Un Estado debe poner suficiente énfasis en promover los valores que pretende formar y unir a sus futuros ciudadanos. Para ello, desde 1993, el Estado mexicano, a través de la Secretaría de Educación Pública (SEP), impulsó la asignatura de Educación Cívica para primaria y Formación Cívica y Ética para secundaria. Estas materias habían desaparecido en los años ochenta cuando se agruparon en la asignatura de Ciencias Sociales y los contenidos de ésta en Historia de México, Historia Universal y Geografía.

En el programa de 1993 para primaria, la Educación Cívica era comprendida como “el proceso a través del cual se promueve el conocimiento y la comprensión del conjunto de normas que regulan la vida social y la formación de valores y actitudes que permiten al individuo integrarse a la sociedad y participar en su mejoramiento” (SEP, 1993a: 123).

Este programa hacía énfasis en los procesos de transformaciones políticas, económicas y sociales que se vivían en país y en la promoción de los derechos humanos, la democracia, el Estado de derecho y la pluralidad política, asimismo, en las organizaciones y los mecanismos de participación ciudadana. Por otro lado, en promover valores y actitudes que dotaran a los estudiantes de bases firmes para formar ciudadanos que conocieran sus derechos y los de los demás, ser responsables en el cumplimiento de sus obligaciones, libres, cooperativos y tolerantes.

Una de las preocupaciones que se tenía en México era la integración de estructuras culturales al orden de la globalización. Para ello, el *currículum* escolar buscaba fortalecer los valores, principios y tradiciones que caracterizaban al país, siendo respetuosos de la diversidad cultural de la humanidad.

En el caso de la educación secundaria, la asignatura de Formación Cívica y Ética se impartía en los dos primeros grados escolares teniendo como referentes los seis años anteriores de educación primaria y los cursos iniciales de la asignatura de Historia.

El propósito general de los cursos de Formación Cívica y Ética era “ofrecer a los alumnos [...] las bases de información y orientación sobre sus derechos y sus responsabilidades, relacionados con su condición actual de adolescentes y también con su futura actuación ciudadana” (SEP, 1993b:127). A su vez, se perseguía que los estudiantes hicieran propios valores sociales como la legalidad, el respeto, la libertad y las responsabilidades personales, la tolerancia, la igualdad de las personas ante las leyes y la democracia como forma de gobierno.

Asimismo, se hacía énfasis en el conocimiento de la organización política e institucional del Estado mexicano tales como el carácter federal de la República, la división de poderes, la relevancia del fortalecimiento de los municipios, las formas democráticas y representativas para la elección de los órganos de autoridad, la identidad y la soberanía nacionales. También se otorgó importancia a que los estudiantes conocieran los principios rectores de las relaciones de México con la comunidad internacional, con el objetivo de vincular la identidad y soberanía nacionales como base de los vínculos con otros países y culturas.

Las asignaturas buscaban otorgar a los estudiantes cuatro líneas fundamentales de aprendizaje: 1) formación de valores, 2) conocimiento y comprensión de los derechos y deberes, 3) conocimiento de las instituciones y de los rasgos principales que caracterizan la organización política de México y; 4) fortalecimiento de la identidad nacional.

Se contaba con algunos problemas políticos, institucionales y de prácticas docentes que dificultan la apropiación de los contenidos de la asignatura en los dos niveles escolares. Por una parte, la Secretaría de Educación Pública y la Comisión Nacional de Libros de Texto Gratuitos (CONALITEG) no dotaron de un texto oficial a los y las estudiantes. Existían algunos textos de editoriales privadas que en algunas ocasiones se apegaban al programa de estudios. Sin embargo, escuelas religiosas editaron sus propios libros y eran utilizados y evaluados en el aula como asignatura del Estado. De igual manera ocurrió con los precios de los textos de editoriales privadas que muchos padres de familia no podían comprar o en algunas zonas del país no se distribuían.

Por otro lado, no se tiene mucha congruencia entre los trabajos de la SEP con la edición de textos gratuitos. Desde el 31 de marzo de 2008 entró en vigor el Programa Integral de Formación Cívica y Ética para educación primaria guiado por competencias escolares pero los contenidos de los textos editados por el gobierno mexicano no corresponden a los contenidos del programa.

Por otra parte, en México es conocido que las asignaturas a las que se les pone mayor atención son Español y Matemáticas debido a que son las materias evaluadas internacionalmente y las que al gobierno le conviene se hagan públicos los resultados, además de ser las materias con las que se asignan mayores recursos económicos tanto a escuelas como profesores dependiendo de las puntuaciones en las pruebas. Las asignaturas de las Ciencias Sociales (Historia, Geografía y Educación Cívica o Formación Cívica y Ética) son aquellas que casi no se abordan en los salones de clase y mucho menos se les otorga mayor importancia por parte de algunos docentes.

Por otra parte, la calidad de la educación y el logro académico no eran claras ni transparentes en México. La Dirección General de Evaluación de Políticas de la SEP era la encargada de construir los exámenes para identificar el logro de los estudiantes en las asignaturas y de dar a conocer los resultados a la población en general. Esto último ocurría dependiendo de las necesidades políticas que se tuvieran que exhibir u ocultar. La SEP se convirtió en juez y parte de la evaluación educativa en México. Muchos años pasaron para que en el país los sistemas de evaluación entraran en un proceso de institucionalización, asociados a promover la calidad educativa.

En 2002, por decreto del entonces presidente, Vicente Fox Quesada, se creó el Instituto Nacional para la Evaluación de la Educación (INEE), cuya misión era la realización de evaluaciones de la calidad del sistema educativo nacional y de los factores que la determinan, así como de la difusión de sus resultados para apoyar la toma de decisiones, la mejora pedagógica en las escuelas y la rendición de cuentas.

Los Exámenes de la Calidad de Logro Educativo (Excale) son pruebas colegiadas, criterios y ligadas al currículum vigente. Estas pruebas son coordinadas por la Dirección de Pruebas y Medición del propio Instituto y elaboradas a través de diferentes comités de investigadores, académicos de diferentes universidades, profesores frente a grupo y representantes de la SEP y su objetivo es dar a conocer la calidad y el logro educativos del sistema educativo nacional en su conjunto, por lo que no se da *ranking* por alumno o escuela.

En las asignaturas de Educación Cívica se han elaborado dos pruebas: en 2006, Excale para 3° de primaria y en 2007, Excale para 3° de secundaria. La prueba para 6° comenzó a elaborarse en febrero de 2008 y hasta el momento sólo se ha realizado un piloteo del total de los reactivos para observar su comportamiento estadístico y seleccionar los mejores.

Sin embargo, los resultados de las pruebas que se han realizado y los elementos que se tiene hasta el momento del piloteo no son alentadores en ninguna de las asignaturas de Ciencias Sociales (Geografía e Historia) y Educación y Formación Cívica y Ética no son la excepción.

2. ¿CÓMO SE CONSTRUYERON LAS PRUEBAS DE EXCALE DE EDUCACIÓN CÍVICA Y FORMACIÓN CÍVICA Y ÉTICA?

El Instituto Nacional para la Evaluación de la Educación (INEE) establece que el objetivo fundamental de la evaluación del aprendizaje es proporcionar un conocimiento general del rendimiento académico de los estudiantes a niveles estatal y nacional (con distintos grados de desagregación por estrato y modalidad educativos), así como de los factores más importantes que explican las diferencias en el logro de los aprendizajes. Este enfoque responde a la necesidad de conocer no tan sólo los resultados académicos de los estudiantes mexicanos, sino también los contextos o entornos inmediatos en los que ocurre el aprendizaje.

Para lograr esta meta el INEE está construyendo una nueva generación de pruebas nacionales: los Exámenes de la Calidad y el Logro Educativos (Excale) con el propósito de evaluar los aprendizajes que logran los estudiantes del Sistema Educativo Nacional (SEN) en sus niveles básicos y medio superior. Para tal efecto, el Consejo Técnico del INEE ha definido que sus pruebas de gran escala tengan tres características distintivas: a) ser de tipo criterial, b) estar alineadas al currículo nacional, y c) tener un diseño matricial. Esto significa que:

- Se diseñan para evaluar dominios más específicos de los estudiantes.
- En México, la educación del nivel básico se ha caracterizado por tener un currículo único. El hecho de contar con un currículo nacional, libros de textos gratuitos y con una formación magisterial uniforme, presenta una gran oportunidad para conocer la calidad de los servicios educativos que ofrece la nación (en términos de lo que logran aprender los estudiantes en relación a lo que se espera de ellos). Esta condición privilegiada no la tienen otros países (como Estados Unidos), quienes para poder evaluar la educación a nivel nacional tienen que darse a la tarea de generar estándares de ejecución en cada disciplina, a fin de alinear con ellos el diseño de las pruebas. Por esta razón, se consideró pertinente que los Excale que el INEE elabora estuvieran alineados al currículo nacional; es decir que su propósito fuera evaluar los aprendizajes pretendidos por el SEN.

El diseño de cada Excale sigue un proceso independiente de desarrollo conforme al orden en que se aplicará. Sin embargo, todos deben compartir el método de construcción y las premisas fundamentales que aseguren un producto de calidad. En el documento Marco Conceptual de las Pruebas Nacionales (INEE, 2004) se explican con cierto detalle los aspectos que se deberán tomar en cuenta para diseñar y construir de manera uniforme los diferentes Excale.

Es importante destacar que para poder diseñar instrumentos nacionales de calidad se requiere el trabajo colegiado de expertos, mismo que se fundamente en un método sólido y consistente con los propósitos que se persiguen. En este punto, cabe mencionar que en el desarrollo de los Excale intervienen diversos especialistas internos y externos, los cuales se agrupan en cinco comités: 1) Académico, 2) Elaborador de Especificaciones, 3) Constructor de Reactivos, 4) Validación y Sesgo, y 5) Establecimiento de Niveles de Logro. Cada uno de éstos cumple con una función específica y complementaria en el proceso de construcción, por lo que su trabajo se programa en forma escalonada y sus productos se convierten en insumos de las siguientes etapas.

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

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

3. RESULTADOS DE LAS PRUEBAS DE EDUCACIÓN CÍVICA DE 3° DE PRIMARIA

Los tres propósitos de la asignatura de Educación Cívica en 3° de primaria son: 1) identificar la estructura de la organización política municipal y estatal, así como los rasgos generales de la organización federal y, reconozcan los derechos y las obligaciones comunes de todos los mexicanos; 2) fortalecer la identidad con valores propios de las regiones y entidades a partir del reconocimiento de sus recursos naturales, su diversidad cultural y social y; 3) identificar valores y símbolos que nos identifican como miembros del Estado mexicano.

En educación primaria la cobertura neta es superior al 94% del grupo en edad escolar (6 a 11 años), la eficiencia terminal se estima en 91.8%, y la absorción de egresados de primaria en la secundaria asciende a 94.9%. Asimismo, el 93.5% de la matrícula de este nivel, se ubica en la modalidad general, el 5.7% en primarias indígenas, y el 0.8% restante en primarias comunitarias.

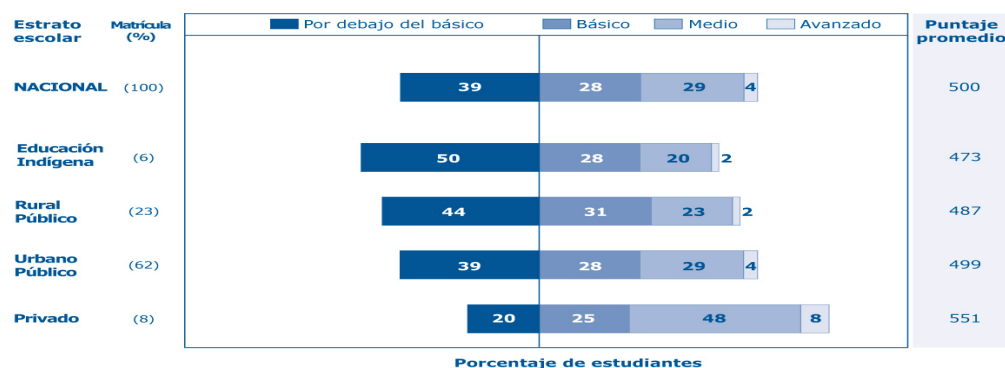
La prueba fue aplicada a 12. 413 estudiantes a nivel nacional, de los cuales 6.245 fueron hombres (50.3%) y 6.168 mujeres (49.7%). Los porcentajes de las edades de los y las estudiantes al momento de la aplicación de la prueba fueron: ocho años o menos, 29.6%; nueve años, 55.8%; diez años, 10.2% y; once años o más; 4.4%.

El Excale fue aplicado a 5.252 estudiantes de escuelas urbanas públicas (42.3%); 2.869 de escuelas rurales (23.1%); 1.214 de escuelas indígenas (9.8%); 232 de cursos comunitarios (1.9%) y; a 2.846 estudiantes de escuelas privadas (22.9%).

Los resultados encontrados fueron los siguientes¹:

Cuatro de cada 10 estudiantes (39%) a escala nacional se encuentran en el nivel *Por debajo del básico*. Tres de cada 10 estudiantes (28%) se ubica en el nivel *básico*; 29% se sitúa en el nivel medio y cuatro alumnos de 100 se ubican en el nivel *avanzado*. En términos generales, se observa que el 67% de estudiantes en este nivel educativo sólo tiene los conocimientos mínimos sobre los objetivos planteados por las autoridades educativas en México.

GRÁFICO 1. NIVELES DE LOGRO A NIVEL NACIONAL



FUENTE:

Instituto Nacional para la Evaluación de la Educación (2007). *El aprendizaje de tercero de primaria en México*, México: INEE.

No obstante, a nivel nacional el nivel escolar donde se encuentra el mayor número de estudiantes es el por debajo del básico (39%). Esto quiere decir que 4 de cada 10 niños o niñas no cuenta con los conocimientos y herramientas suficientes para continuar aprendiendo en los siguientes grados y niveles cognitivos. Sumado a este porcentaje, 28% se considera con los conocimientos básicos para continuar aprendiendo en la asignatura. Esto quiere decir que 67% de los estudiantes de 3° de primaria en el país no cuenta o sólo tiene los conocimientos mínimos sobre lo trabajado en los salones de clase.

De igual manera es cuestionable que por estratos escolares, la mitad de los estudiantes de educación indígena no cuenta con los elementos para continuar eficientemente en la asignatura y que otro 28% sólo cuenta con los elementos mínimos, es decir, 78% de los y las estudiantes de este sector social no han logrado una apropiación de los conocimientos curriculares propuestos por la SEP. Ante esta situación surgen interrogantes respecto a las características socioeconómicas y culturales que contradicen el aprovechamiento escolar exigido por las autoridades mexicanas en materia educativa. De igual manera ocurre con las escuelas rurales públicas donde la mayoría de sus estudiantes se encuentran en los niveles por debajo del básico y básico.

Se encontró que las escuelas que tuvieron mejores resultados fueron las escuelas privadas donde la mayoría de sus estudiantes se concentraron en el nivel medio. En segundo lugar se encontraron las escuelas urbanas públicas, pero con un aumento de 19% con respecto a las privadas en el nivel por debajo del básico, es decir, el doble.

¹ Los únicos resultados publicados oficialmente por el INEE son los de 3° de primaria

Los aprendizajes que mejor dominan los estudiantes de tercero de primaria, a nivel nacional, son: El sentido de pertenencia a la comunidad, a la Nación y a la humanidad (57%), Conocimiento y comprensión de los derechos y los deberes ciudadanos (50%), Formación de valores y naturaleza humana (49%).

De los conocimientos y habilidades que con menor destreza manejan los alumnos son los relacionados con la *organización política del Estado mexicano* (39%) y con la *democracia y participación ciudadana* (38%).

En términos de aprendizajes más específicos, los resultados obtenidos por los estudiantes de tercero de primaria indican que los alumnos dominan las siguientes habilidades y conocimientos de Educación cívica:

3.1. Sentido de pertenencia a la comunidad, a la Nación y a la humanidad

Siete de cada 10 alumnos pueden identificar los símbolos patrios y reconocer que éstos representan la identidad nacional. Seis de cada 10 identifican manifestaciones culturales del país como son: bailes, alimentos, juguetes, etcétera; pero sólo la mitad identifica las fiestas tradicionales nacionales y la tercera parte las costumbres familiares.

3.2. Conocimiento y comprensión de los derechos y los deberes

Poco más de la mitad de los alumnos reconoce y valora sus derechos en cuanto a protección, educación, salud, respetar y ser respetado, así como los beneficios que éstos conllevan. Cuatro de cada 10 reconocen y valoran su derecho a la recreación y descanso, así como a tener un nombre y una nacionalidad. Siete alumnos de cada 10 identifican los cuidados que hay que tener de los servicios públicos, pero sólo el 44 % reconoce el uso adecuado de éstos.

3.3. Formación de valores y naturaleza humana

En el ámbito familiar, seis de cada 10 alumnos reconocen la importancia de la colaboración, la solidaridad, el beneficio de la comunicación y el diálogo, así como la necesidad de distribuir tareas y responsabilidades entre los miembros de la familia; además, reconocen el valor de satisfacer las necesidades sin aprovecharse de los demás. En el ámbito local, sólo cuatro de cada 10 alumnos identifican al trabajo como la única manera digna de obtener satisfactores; valoran la interdependencia laboral a favor del bien común e identifican la importancia y las ventajas que trae consigo el trabajo en grupo para resolver problemas de la localidad.

3.4. Organización política del Estado mexicano

Dos terceras partes de los alumnos identifican al Presidente de México, sin embargo, a pesar de que el currículo está enfocado a los aspectos locales de la entidad, sólo una tercera parte reconoce e identifica a las autoridades locales y nacionales. Poco más de la mitad de los alumnos identifica las funciones del poder judicial de la entidad, pero sólo dos de cada 10 identifican las funciones del Poder Legislativo y del Ejecutivo de la entidad.

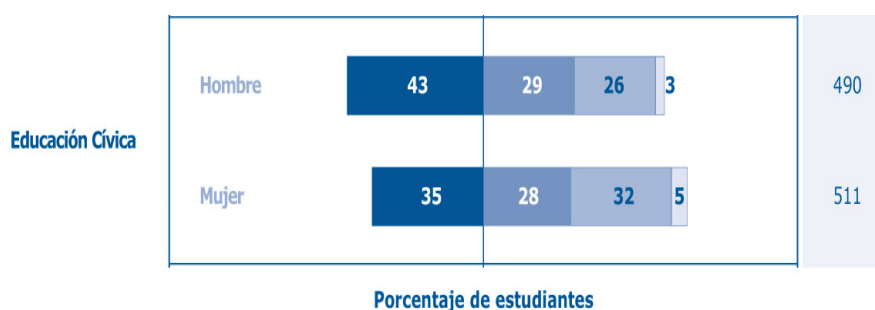
Poco más de la mitad de los alumnos puede escribir el nombre o sigla de algún partido político registrado ante el IFE, pero sólo la tercera parte identifica las organizaciones sociales locales. Cuatro de cada 10 alumnos identifican que la integración del ayuntamiento es el resultado de un proceso de elección; asimismo, reconocen al federalismo por medio de la identificación de autoridades, leyes o normas;

además, reconocen los beneficios sociales y políticos que conlleva contar con la Constitución, pero sólo tres de cada 10 alumnos reconocen que la Constitución es la Ley Suprema de todo el país.

3.5. Democracia y participación ciudadana

Poco más de la mitad de los alumnos identifica las organizaciones sociales y políticas, reconoce los fines de éstas e identifica la importancia del trabajo en grupo para resolver problemas de la localidad. Tres alumnos de cada 10 reconocen las reglas de convivencia de las organizaciones sociales y sólo la cuarta parte reconoce la importancia de ellas a nivel local. Poco más de la tercera parte de los alumnos identifica los beneficios del derecho a elegir y ser electo e identifica la edad en que se adquiere este derecho, así como elegir al gobierno de la entidad e identificar los periodos de elección federal y estatal.

GRÁFICO 2. NIVELES DE LOGRO POR SEXO

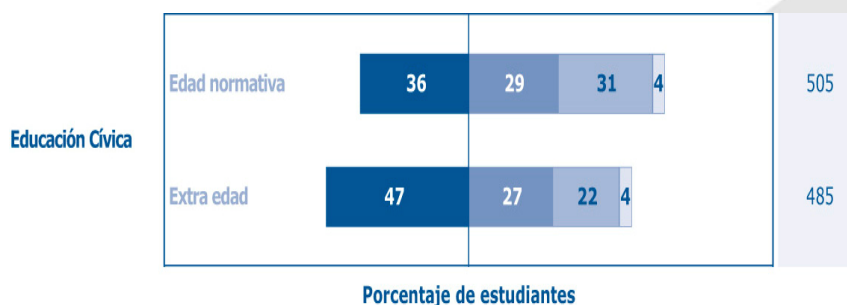


FUENTE: Instituto Nacional para la Evaluación de la Educación (2007). *El aprendizaje de tercero de primaria en México*, México: INEE.

Asimismo, se pudo observar que las mujeres tuvieron mejor desempeño que los hombres, con un puntaje mayor de 21 puntos. Ellas lograron ubicarse de manera semejante en el nivel por debajo del básico que en el medio. A su vez, ellas tuvieron el mayor número de estudiantes en el nivel avanzado. Por los hombres, 4 de cada 10 tienen un nivel de logro por debajo del satisfactorio para continuar aprendiendo en la asignatura. Tres de ellos sólo cuentan con el nivel mínimo aceptable.

Cabe destacar que los mayores porcentajes tanto de hombres como de mujeres se encuentran en el nivel por debajo del básico, lo cual conmina a poner atención en la formación ciudadana y ética en los dos sexos.

GRÁFICO 3. NIVELES DE LOGRO POR EDADES NORMATIVA Y EXTRAEDAD



FUENTE: Instituto Nacional para la Evaluación de la Educación (2007). *El aprendizaje de tercero de primaria en México*, México: INEE.

Por otro lado, se puede observar que casi la mitad de los estudiantes que tienen más de nueve años y que cursan el 3° de primaria se encuentran en una situación donde no podrán continuar aprendiendo en los años siguientes porque no cuentan con los conocimientos básicos para seguir con el proceso de aprendizaje. Debemos señalar que, en este caso, 74% de los niños y niñas tienen un nivel por debajo y básico en este grado escolar, es decir, los niveles medio y avanzado representan sólo una cuarta parte de la muestra.

Las niñas y los niños en edad normativa obtuvieron 20 puntos más que los que están en situación de extra edad. Ello se ve reflejado en la disminución de estudiantes en el nivel por debajo del básico y el aumento considerable en el nivel medio. Estos datos manifiestan las acciones necesarias que se deben tomar en cuenta para evaluar con carácter normativo en el sistema educativo nacional y evitar la reprobación ya que sus consecuencias se manifestarán de pero manera en los estudiantes de extraedad.

No obstante, 65% de los y las estudiantes se encuentran en los niveles por debajo del básico y básico, lo cual se traduce en que la mayoría de estos estudiantes sólo tienen un nivel mínimo de conocimientos en esta materia y en su mayoría no cuenta con estos elementos. Recordemos que tanto en los estudiantes de edad normativa como en situación de extraedad el mayor porcentaje de los cuatro niveles fue ubicado en el nivel por debajo del básico.

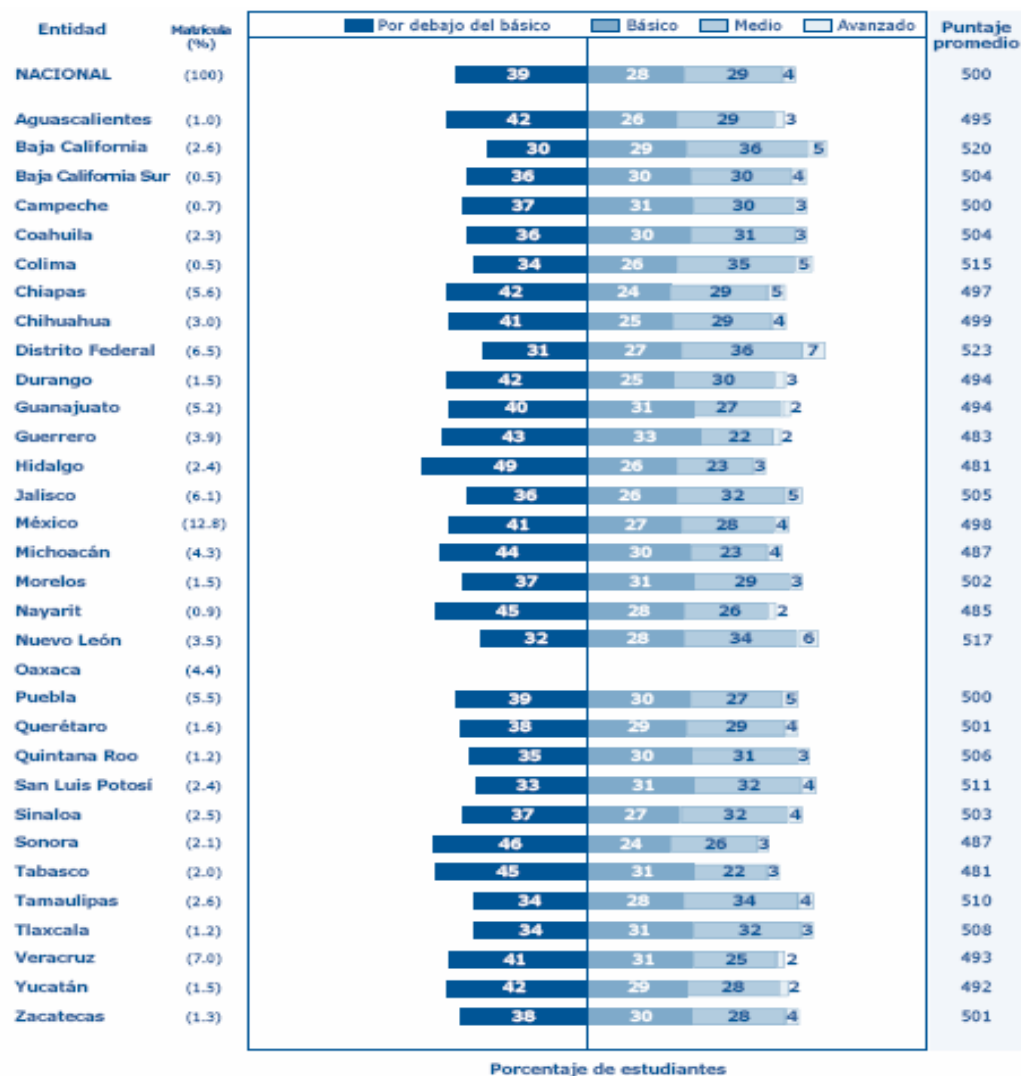
Por otra parte, los estados de la república donde se encontraron los mayores niveles de logro en Educación cívica fueron el Distrito Federal, Baja California y Nuevo León con los puntajes promedio más altos del país. Los estados que tuvieron el peor nivel de logro en esta asignatura fueron Tabasco, Hidalgo y Guerrero.

El Índice de Desarrollo Humano (IDH) en 2006 señalaba que tanto el Distrito Federal, Nuevo León y Baja California ocupaban los tres primeros lugares en su clasificación con índices de 0.8837, 0.8513 y 0.8435 respectivamente. Esto quiere decir que las condiciones de vida de estas entidades son semejantes a las de algunos países como República Checa, Barbados, Eslovaquia, Uruguay, Costa Rica y los Emiratos Árabes Unidos. Este mismo documento determinó que eran las entidades con mayor índice de salud, educación y niveles de ingreso en el país, lo cual nos invita a pensar que las características socioeconómicas y culturales son concluyentes en los niveles de logro educativo.

Asimismo, Tabasco ocupó el lugar 21, Hidalgo el 27 y Guerrero el 30 en el IDH con 0.7800, 0.7645 y 0.7390 respectivamente. Estos índices son semejantes a países como Tailandia, Samoa, Ecuador, Filipinas, Palestina y Azerbaián. Esto nos indica que el mandato constitucional y las políticas públicas en materia educativa no han dado resultados y los estados que cuentan con mayores recursos son los que obtienen mayores niveles de logro educativo, lo cual está ocasionando una enorme brecha en cuanto al conocimiento de los contenidos de la asignatura de Educación cívica.

Se debe destacar que el puntaje medio de la medición fue de 500 puntos y 17 de las 32 entidades federativas tuvieron un puntaje mayor a la media. Sin embargo, los reactivos con dificultad más alta correspondieron a 771 puntos.

GRÁFICO 4. NIVELES DE LOGRO POR ENTIDAD FEDERATIVA



FUENTE: Instituto Nacional para la Evaluación de la Educación (2007). El aprendizaje de tercero de primaria en México, México: INEE.

4. APROXIMACIONES DE LOS RESULTADOS DE LAS PRUEBAS DE FORMACIÓN CÍVICA Y ÉTICA DE 3° DE SECUNDARIA²

El nivel secundaria brinda servicio a 6 millones de estudiantes, cifra equivalente al 88 por ciento del grupo de edad entre 12 y 14 años. Sin embargo, la eficiencia terminal es del orden de 79%, lo que significa que al menos 400 mil estudiantes, de los más de 2 millones que ingresan cada año, no concluyen sus estudios en el tiempo previsto.

² Los resultados oficiales no se han presentado a la fecha en que se escribió este artículo, por lo cual sólo se darán a conocer aproximaciones del estudio.

La prueba de Formación Cívica y Ética para 3° de secundaria se aplicó en el año 2007 a 50 932 estudiantes, de los cuales 24 037 fueron hombres y 26 895 mujeres de escuelas generales, técnicas, privadas y telesecundarias.

Cabe señalar que al término de este grado escolar, los y las estudiantes estarían a tres años de ejercer de manera legal en el país y para ello se debía contar con un amplio criterio y sentido de responsabilidad social y política.

TABLA 1. NÚMERO DE ALUMNOS POR SEXO Y EDAD

Sexo del alumno * Edad del alumno en años									
Modalidad educativa o estrato			Edad del alumno en años						Total
			13	14	15	16	17	18	
Generales	Sexo del alumno	Hombre	16	2615	3903	813	169	26	7542
		Mujer	19	3460	4348	636	139	26	8628
	Total		35	6075	8251	1449	308	52	16170
Privadas	Sexo del alumno	Hombre	9	1569	2491	405	48	5	4527
		Mujer	8	1951	2781	302	24	5	5071
	Total		17	3520	5272	707	72	10	9598
Técnicas	Sexo del alumno	Hombre	15	2504	3768	768	189	28	7272
		Mujer	27	3217	4364	638	146	19	8411
	Total		42	5721	8132	1406	335	47	15683
Tele secundarias	Sexo del alumno	Hombre	8	1245	2238	824	286	95	4696
		Mujer	9	1473	2422	616	195	70	4785
	Total		17	2718	4660	1440	481	165	9481

FUENTE: Base de datos del Instituto Nacional para la Evaluación de la Educación

Siguiendo los niveles de logro propuestos por el INEE se obtuvieron los siguientes resultados:

El nivel avanzado, tuvo un porcentaje cercano al 2% de los estudiantes de secundaria. Este nivel demandó de los alumnos, en términos cognoscitivos, que distinguieran de manera "sutil" entre valores que si bien conceptualmente saben que son distintos, sólo al reflexionar sobre ellos se puede establecer que existen elementos que comparten e incluso que pueden traslaparse; es decir, implica más que su mera identificación, ya que era necesario su comprensión y aplicación de cada uno de los valores que se presentan como opciones de respuesta.

Bajo estos parámetros el nivel de logro avanzado se caracteriza por que las alumnas y los alumnos de este nivel presentan un dominio muy elevado de conocimientos, habilidades y destrezas que les permitan *reflexionar* y *valorar* sobre las diversas situaciones que se les presentan en los reactivos en dos dimensiones tales como ética y cívica y no sólo identificar, comprender y/o aplicar en una situación los conocimientos adquiridos, como en los niveles de logro previos.

En relación con la dimensión ética, los alumnos *reflexionan* sobre la importancia del ejercicio de una sexualidad responsable, la importancia de la elaboración de un proyecto de vida personal que considere las peculiaridades del entorno social, la responsabilidad de su actuar que tienen para consigo mismos y con el entorno social (familia, escuela, trabajo y comunidad) y sobre las prácticas familiares cotidianas socialmente aceptadas y que constituyen estereotipos de género.

Con respecto a la dimensión cívica, los alumnos *comprenden* las diferentes funciones de los poderes del Estado democrático, *diferencian* entre los valores que se aplican en la Constitución Política de los Estados Unidos Mexicanos e *identifican* el juicio de amparo o queja ante los organismos públicos como mecanismos para defender los derechos humanos y las libertades constitucionales.

Temas como libertad, igualdad, equidad, justicia, tolerancia, diálogo, reconocimiento de los beneficios de la separación de poderes, las garantías individuales que otorga la constitución de 1917 fueron contenidos que tuvieron un porcentaje de respuesta de 31%, al igual que la identificación de estereotipos de género y el reconocimiento de un proyecto de vida. Asimismo los reactivos que tuvieron un grado de dificultad mayor a 800³ y que fueron contestados por los estudiantes ubicados en este nivel fueron tres. El primero hacía referencia al reconocimiento de roles que se le dan a mujeres y hombres en algunas situaciones y tuvo un nivel de dificultad de 811.14 con 19% de respuesta. Con ese mismo porcentaje de respuesta se encontró a la democracia como forma de organización social, pero para los estudiantes tuvo mayor dificultad con 818.25. El reactivo que implicó mayor dificultad para los estudiantes correspondió a las ventajas de hacer proyectos de vida durante la juventud con 16% de respuesta y un nivel de dificultad de 850.35.

Con base en las demandas cognoscitivas anteriormente descritas, muy pocos alumnos de tercer grado de secundaria a nivel nacional pudieron alcanzar este nivel de logro educativo. Esto implica que sólo 2 de cada 100 aspiran a este nivel. Este porcentaje minoritario muestra que existe una gran cantidad de variables escolares que podrían dificultar el desempeño en la asignatura como la práctica memorística para la adquisición de los valores, la falta de contextualización en el ejercicio de los mismos y la falta de congruencia entre el discurso que fomenta la asignatura con la realidad cotidiana de las instituciones.

A pesar de la realidad educativa existen alumnos que cuentan con elementos extraescolares que favorecen su desempeño; tales como, variables de contexto que lo fortalecen y un capital cultural amplio. Estos factores desafortunadamente a nivel nacional sólo se reflejan en un porcentaje muy reducido de la población mexicana, por lo que sólo algunos estudiantes aspiran a niveles reflexivos de gran alcance como el nivel de logro avanzado en la asignatura de Formación Cívica y Ética.

Asimismo, el nivel de logro medio obtuvo un porcentaje cercano a una quinta parte del total de estudiantes. Este nivel se caracterizó porque las alumnas y los alumnos presentan un dominio adecuado de conocimientos, habilidades y destrezas que les permitan reconocer, comprender y/o aplicar en una situación los conocimientos adquiridos en las dos dimensiones: cívica y ética.

Con respecto a la primera, se demanda el *reconocimiento* de la influencia de la participación ciudadana, de las acciones de las organizaciones sociales democráticas, los valores para *establecer relaciones* democráticas (diálogo, respeto, tolerancia, libertad) y sociales (familia, escuela, trabajo y sociedad), el ejercicio de los derechos y obligaciones de los ciudadanos y las *relaciones* entre las garantías constitucionales y los derechos humanos. En relación con la segunda, los alumnos *reconocen* la dimensión afectiva y la dignidad de las otras personas en el ejercicio de la sexualidad, así mismo *comprenden* que una vida sana implica un equilibrio entre el bienestar físico, mental y social.

Los niveles del porcentaje de respuesta se ubican entre 34% y 53%. El reactivo último al punto de corte al nivel avanzado tuvo un nivel de dificultad de 690.96 y un porcentaje de respuesta de 34% y el

³ La escala manejada en INEE es de 200 a 800.

contenido que se abordaba era el concepto de adicción. De igual manera el más fácil, dentro de este rango, tuvo una dificultad de 579.01 y 52% de respuesta y el tema que se abordó fue precisar el concepto de valor desde el punto de vista de la convivencia humana.

Con base en las demandas cognoscitivas anteriormente descritas, sólo alrededor de una quinta parte de los alumnos de tercer grado de secundaria a nivel nacional están en posibilidades de alcanzar este nivel de logro educativo.

Este porcentaje refleja la interrelación de varios factores que dificultan que la mayoría de los alumnos adquieran un nivel de comprensión y aplicación sobre los conocimientos y las habilidades que se fomentan en el aula. Algunos de estos factores son la práctica memorística para la adquisición de los valores, la falta de contextualización en el ejercicio de los mismos y la falta de congruencia entre el discurso que fomenta la asignatura con la realidad cotidiana de las instituciones educativas.

En el nivel básico se ubicaron más de la mitad de los alumnos. En términos prácticos, este dato se traduciría que en un grupo de 40 alumnos, 24 tienen los conocimientos suficientes para continuar de manera satisfactoria sus estudios en la asignatura.

Consideramos que el mayor porcentaje de alumnos se encuentra en este nivel debido a que en los últimos años la participación ciudadana ha enriquecido los conceptos de organización, democracia, Estado democrático entre otros. Muchos reactivos dentro de este rango cuestionaron sobre el deber ser y los alumnos pueden identificarlos e interpretar este tipo de respuestas de carácter ético.

Por otra parte, los alumnos de nivel básico tienen un dominio esencial de conocimientos mínimos que les permiten interpretar situaciones. Si bien es cierto que los alumnos de este nivel presentaron carencias reales en cuanto al manejo de los conocimientos previos y a la transversalidad curricular, este nivel refleja a aquellos que tienen un dominio suficiente de los conceptos de la asignatura.

Contenidos como la importancia de la Constitución, reconocimiento de artículos constitucionales, participación ciudadana, derechos y responsabilidades en la familia, cooperación, amistad, trabajo digno, responsabilidad social, prevención de adicciones, medio ambiente, reglas sociales, violencia intrafamiliar, identidad y el propio reconocimiento de la asignatura en su vida cotidiana se encuentran en esta esfera de contenidos.

El Comité que estableció los niveles de corte⁴ observó que la situación económica a esas edades, también incide en el resultado, ya que repercute en la asistencia, aprovechamiento y rendimiento escolar, los alumnos tienen la potencialidad y no la desarrollan porque no tienen recursos suficientes e incluso ellos participan en el ingreso familiar. Asimismo, existe una enorme contradicción entre la realidad social y los preceptos de la asignatura. La educación secundaria es formativa, pero la sociedad en su conjunto la ha limitado por su praxis.

Por otra parte, los estudiantes ubicados en el nivel por debajo del básico identificaron los actores a los que hacen referencia los artículos 3° y 31 constitucionales (81%); identifican los riesgos en la salud que tiene la práctica sexual sin protección (83%) al igual que el concepto de droga (84%) y reconocen las

⁴ Este comité nombrado niveles de logro se divide en dos. El primero es el Comité Académico, aquel que diseña la prueba. El segundo es el Comité de niveles de corte y participan profesores frente a grupo de escuelas generales, técnicas, privadas, telesecundarias.

ventajas personales y sociales que tiene el uso de anticonceptivos (91%). Los niveles de dificultad para este nivel son de 242.06 a 371.99.

En términos prácticos se estima que en un grupo de 40 alumnos, aproximadamente siete tienen limitaciones reales para dominar contenidos esenciales de su formación no sólo como estudiantes sino como individuos, lo que les impide continuar sus estudios de manera satisfactoria. De igual manera, sus respuestas no aseguran (a excepción de los artículos constitucionales) que dicha información haya sido aprendida en el aula, ya que en los medios de comunicación es frecuente encontrarse con ella.

5. CONCLUSIONES

La educación en cuanto a la formación de ciudadanos y valores para la cohesión social no ha sido, desde una política educativa del Estado mexicano, prioridad. Podemos observar en los resultados de la evaluación a nivel nacional que el sistema educativo en su conjunto no pudo elevar el nivel de logro académico en las asignaturas de Educación Cívica y Formación Cívica y Ética. Ello lo podemos comprobar en los porcentajes de las dos pruebas.

En 3° grado de educación secundaria se logró disminuir a cerca de 15% el número de estudiantes ubicados en el nivel por debajo del básico, respecto al 39% de 3° grado de educación primaria. Sin embargo, el nivel básico muestra un fuerte aumento en los conocimientos mínimos sobre la asignatura en educación secundaria respecto a su equivalente en la primaria. Se pasó de 28% a más de 60% de los estudiantes que sólo tienen un nivel de conocimientos elementales sobre cuestiones cívicas y éticas. Esto es aún más alarmante si tomamos en cuenta que se abordó durante nueve años y que será lo último de manera obligatoria que el Estado mexicano promueva entre sus jóvenes, ya que de acuerdo con los resultados del *Panorama de la educación 2008*, elaborado por la Organización de Cooperación y el Desarrollo Económico (OCDE), México tiene una de las tasas más bajas de cobertura educativa entre los jóvenes de 15 a 19 años, pues 45 por ciento no asiste a la escuela. De éstos, 62 por ciento tiene empleo y el 38 por ciento restante no estudia ni trabaja, lo que representa uno de los desafíos más grandes para el sistema nacional de enseñanza.

Por otro lado, si sumamos el nivel por debajo del básico y básico para poder identificar aquellos estudiantes de menor nivel de logro educativo respecto a estas asignaturas tenemos un aumento de manera negativa en los porcentajes. Los resultados de Excale para 3° de primaria nos dieron una suma de 67%, que de por sí es alarmante, ya que más de la mitad de los estudiantes de este nivel tienen un nivel muy bajo de logro escolar, en 3° de secundaria se incrementó a casi el 80% de estudiantes en estos niveles. Esto quiere decir que a nivel 3° de primaria siete de cada diez alumnos tiene un nivel mínimo suficiente de aprendizaje de la materia, mientras que en educación secundaria, al concluir la educación básica obligatoria, 8 de cada 10 alumnos presentan la misma situación.

En México la edad para ejercer el voto y participar de manera legal en la vida política del país es de 18 años, es decir, tres años después de egresar de la educación media básica. En México, los habitantes de 15 años y más, en promedio tienen 8.1 grados de escolaridad, lo que significa un poco más del segundo año de secundaria, es decir, los conocimientos adquiridos en materia cívica y ética los adquirieron en estos años de instrucción. Si el nivel empeora con los años de educación, esto quiere decir que el nivel de conocimiento al egresar de la educación básica e insertarse a un siguiente nivel o la vida laboral tendrá

menos herramientas en cuanto al conocimiento de sus derechos ciudadanos, la organización del Estado mexicano, identidad nacional y valores para convivencia.

Por otra parte, podemos identificar que existe un gran rezago en el nivel de logro en Educación Cívica en las escuelas indígenas donde se encuentra el 50% de su población en el nivel por debajo del básico y 28% en el básico. Esto se traduce en que casi el 80% de los estudiantes de escuelas indígenas tienen un nivel muy deficiente en los conocimientos de la asignatura. El porcentaje de esta educación es muy semejante, en general, al de toda la educación secundaria. Sin embargo, en las escuelas indígenas hay un mayor porcentaje de estudiantes en el nivel más bajo. Otro foco rojo, lo encontramos en las escuelas rurales públicas donde el nivel por debajo del básico representa 44% de sus estudiantes y en el nivel siguiente 31%, es decir, la educación cívica en esas escuelas tiene a tres cuartas partes de sus alumnos con un nivel mínimo de logro educativo.

Las escuelas ciudadinas públicas tampoco son la excepción, ya que cuentan con 39% de estudiantes en el nivel más bajo y 28% en el nivel básico, al igual que las escuelas indígenas. Esto se traduce en que siete de cada diez estudiantes de primarias urbanas públicas sólo tienen las herramientas mínimas para poder seguir aprendiendo. Las escuelas privadas, al contrario, cuentan con 56% de sus estudiantes en los niveles medio y avanzado (en este último ocho de cada 100, el doble de las escuelas urbanas públicas y seis veces mayor que en las indígenas y rurales). De igual manera, las escuelas privadas fueron las únicas que estuvieron por arriba del promedio con 551 puntos, mientras que las urbanas públicas obtuvieron 499, las rurales públicas 487 y las indígenas 473 puntos.

Lo anterior demuestra como el gobierno mexicano, a través de la Secretaría de Educación Pública, no ha logrado establecer una serie de políticas educativas que fortalezcan el mandato constitucional de otorgar educación de manera obligatoria a todos los niños, niñas y jóvenes que se encuentran en edades de escolaridades básicas. En cambio, el mayor porcentaje de niveles de logro educativo en las escuelas privadas refleja como el gobierno de este país ha cedido parte de sus obligaciones a los particulares.

Los diferentes comités de especialistas que han participado en la elaboración de las pruebas han coincidido que existen grandes contradicciones en los contenidos creados por la SEP y las diversas realidades donde se desenvuelven de manera cotidiana los y las estudiantes, además de que no existe una buena capacitación a los profesores para impartir las materias. Esto es otro de los errores programáticos en la formación de los docentes, ya que cuando éstos egresan de las Escuelas Normales los planes y programas han cambiado y los cursos de actualización son deficientes. Asimismo, considerar asignaturas como más importantes que otras y con base en ello otorgar mayores recursos y reconocimientos al propio Estado o gratificaciones económicas a los maestros han ocasionados una fuerte contradicción ética que impide a los estudiantes tener una educación integral y adoptar y fortalecer herramientas que les permitirán construir seres humanos con capacidades de discernimiento moral, ético y ciudadano.

Se debe tomar en cuenta que en México en 2002, según el Instituto Nacional de Estadística, Geografía e Informática (INEGI), 3.3 millones de niños y niñas de entre 6 y 14 años realizaban algún trabajo, con una distribución de 50% mujeres y 50% hombres. Las mujeres realizan mayor trabajo del hogar (no remunerado) mientras que los hombres mayor trabajo económico. Esto debido a que 51% trabajan en un negocio familiar por falta de recursos económicos para contratar un empleado, 27% laboran para comprarse cosas que sus padres no pueden y, 22 % lo hace porque en su familia necesita ese ingreso. En las zonas rurales e indígenas, se tienen 48.2% de niños y niñas en trabajos agropecuarios, 14.2% se

dedican al comercio, 20.5% son artesanos o ayudantes, 11% empleados y 4.3% son trabajadores domésticos. De ellos y ellas, sólo 28% recibe algún pago y 72% no lo recibe. Esto se refleja en el porcentaje de niños, niñas y jóvenes que estudian. El porcentaje de las personas de entre 6 y 11 años de edad (edades de educación primaria) que trabajan y no estudian es de 9%, mientras que los de 12 a 14 años (edades de educación secundaria) la cifra se incrementa hasta 34%. Por lo cual, el Estado debe brindar las mismas oportunidades a los y las estudiantes para que un discurso de valores y oportunidades se fortalezca y no se contradiga como hasta la fecha ha sido.

Son muchos los retos que tiene el Estado mexicano en materia de educación y mayores las exigencias que presenta la sociedad ante fenómenos políticos, económicos, sociales y culturales que están en constante cambio. Sin embargo, descuidar la formación de ciudadanos y ciudadanas con suficientes herramientas políticas, legales, institucionales y axiológicas para actuar en su país implica un enorme riesgo de fracturas y rupturas del lazo social, así como, una deficiente inserción a una comunidad global cada vez más excluyente y más exigente a los seres humanos. Por otra parte, crear sociedades y gobiernos más congruentes con los valores que enseñan fortalecerá la posibilidad de tener menos inequidad económica y social, menores problemas de violencia y adicciones, así como mayores oportunidades de desarrollo personal y colectivo, mejor y más justa distribución de las riquezas y de los conocimientos, mayor respeto y aprecio por la diversidad, mayor estima por la justicia, la libertad, la solidaridad, cooperación y, por tanto, mejores y más dignas posibilidades de vida.

REFERENCIAS BIBLIOGRÁFICAS

- Instituto Nacional para la Evaluación de la Educación (2004). *El aprendizaje en tercero de primaria en México*. México: INEE.
- Instituto Nacional para la Evaluación de la Educación (2004). *Marco Conceptual de las Pruebas Nacionales*. México: INEE.
- Secretaría de Educación Pública (1993a). *Educación básica primaria. Plan y programas de estudio 1993*. México: SEP.
- Secretaría de Educación Pública (1993b). *Educación básica secundaria. Plan y programas de estudio 1993*. México: SEP.



ISSN: 1989-0397

LA EVALUACIÓN DE LA ESCRITURA PARA LA EVALUACIÓN DE LA LECTURA

Fabio Jurado

Revista Iberoamericana de Evaluación Educativa 2009 - Volumen 2, Número 1

<http://www.rinace.net/riee/numeros/vol2-num1/art7.pdf>

Fecha de recepción: 21 de marzo de 2009

Fecha de comunicación de dictamen: 25 de abril de 2009

Fecha de aceptación: 28 de abril de 2009



1. UNOS PROLEGÓMONOS

La escritura y la lectura constituyen dos prácticas decisivas en el desarrollo intelectual y en la formación política de los sujetos. Es un compromiso de la escuela concentrar todas sus fuerzas hacia el acceso y el dominio pleno de estas prácticas que, en general, son un derecho y una vía para cohesionar la ciudadanía y la posibilidad de la democracia. Pero no podemos hablar de lectura y de escritura de manera llana, como si entendiéramos con facilidad lo que es saber leer y saber escribir. Hemos de reconocer que son aprendizajes complejos, que permanecen durante toda la vida, y que no hay una única manera de saber leer y escribir, como lo pretenden los programas de alfabetización que se trazan en nuestros países. Es necesario distinguir entre la escuela que simplemente alfabetiza y la escuela que orienta sus procesos para aprender a leer críticamente los textos que circulan en la vida académica y en la cultura en la que estamos inmersos.

La escuela ha tendido a través de su historia a anclar sus propósitos en la alfabetización inicial, es decir, enseñar las letras para saber combinarlas y formar palabras y luego saber construir oraciones; es lo esencial que se requiere para saber leer avisos publicitarios, titulares de prensa, boletines, cuentas de cobro, citaciones, saber firmar para votar... En esta perspectiva las cartillas y los libros de texto reiteran en la enseñanza de la morfología de las palabras, la identificación de las partes de la oración, el listado de vocabulario "para hablar y escribir bien", los sinónimos y antónimos, las normas ortográficas, las conjugaciones verbales... Es lo propio de la alfabetización.

Felipe Garrido señala cómo el reto de todo proyecto educativo hoy es el de desbordar la mera alfabetización para acentuar más la formación del lector crítico; por eso cree necesario "convertir en lectores a todos los alumnos, profesores, padres de familia...":

Está alfabetizado quien puede simular la lectura de un texto cuando lo tiene enfrente. Digo simular porque esta operación no incluye la comprensión del texto; esta operación supone que la comprensión es una segunda etapa en la adquisición de la capacidad de leer. Y la consecuencia de esa simulación es que avancemos por la vida leyendo –comprendiendo– a medias, aprovechando a medias las lecturas a nuestro alcance.

(Garrido, 1999:66)

Pareciera que luego de la alfabetización hubiese que adelantar grandes campañas para formar lectores, como si el lector se constituyera en una fase posterior y no paralelamente. Leer "a medias" es leer desde una dimensión local del texto (leer palabras o frases) y no desde la coherencia global (cfr. van-Dijk, 1982); leer "a medias" es también leer sin una actitud dispuesta hacia la interpelación del texto, hacia su interrogación y su réplica, lo que presupone una incomunicación con el texto. Cabría preguntarse, en consecuencia, qué es ser un lector, pregunta que se planteaba Pedro Salinas (1955) cuando invitaba a distinguir entre el leedor (el que lee a medias) y el lector (el que lee desde una actitud interrogadora y bucea en la profundidad del texto).

Ese lector que bucea en la profundidad del texto es el lector crítico, un lector que resulta extraño y díscolo para una sociedad que busca homogenizar a todos sus miembros. El lector crítico dialoga con el texto porque al interpretar actualiza el repertorio cognitivo y experiencial (la Enciclopedia del lector, que llamara Eco) y entreteje relaciones diversas que convergen en puntos de vista, en posicionamientos frente al texto y frente al mundo. La lectura es posible por un diálogo con el texto, lo que supone reconocer que

los textos piensan, provocan y seducen hacia la acción. La fuerza dialogante con los textos se evidencia cuando en ellos dejamos las huellas, o improntas, de la travesía interpretativa, a través de los subrayados, de notaciones en los márgenes (signos de interrogación o de exclamación) e inclusive enunciados y remisiones a otros textos. Estas improntas constituyen impulsos hacia la escritura, porque ese diálogo es tan fuerte que el lector experimenta la sensación de estar escribiendo a medida que avanza en la lectura.

El aprendizaje de la lectura presupone el aprendizaje de la escritura, si bien los ritmos para el dominio de una y otra no son equivalentes: la escritura como acto productor de conocimiento se caracteriza por la morosidad. Esta morosidad es natural, pues se trata de proyectar el pensamiento en una representación exterior: la escritura, que resulta siendo una representación de la representación (el pensamiento es representación). La morosidad en el acto de escribir deviene de la dificultad para darle cuerpo discursivo a las ideas, sobre todo porque no hay una equivalencia entre las ideas que procesa el pensamiento y la construcción lingüístico-discursiva: las palabras fungen de mediadoras entre las ideas y su exteriorización.

También sobre la escritura es necesario puntualizar aspectos relacionados con su singularidad y sus efectos potentes en la cualificación del pensamiento. La escritura constituye la mejor estrategia para evaluar los grados de apropiación semántica (la lectura) de los textos, dado que no se escribe desde un lugar cero sino desde múltiples y diversos lugares; estos lugares son los textos aprehendidos por los lectores; no de otro modo se puede leer y se puede escribir: desde ciertos lugares (es el repertorio de cada lector). Esta experiencia, cuando es auténtica, es decir, cuando no es estereotipada o artificial como regularmente ocurre en el contexto escolar (hay que escribir sobre el paseo o hay que escribir un resumen del libro, sólo para obtener una calificación), está asociada con el asombro y con una cierta sensación de misterio y de deseo por descubrir cosas nuevas.

Al comienzo de *Los límites de la interpretación*, Umberto Eco referencia un pasaje del libro de John Wilkins en el que narra cómo un esclavo se come en el camino gran parte de los higos que su amo estaba enviando a un amigo, junto con una carta. En la carta se informa del envío de una cantidad determinada de higos; dicha cantidad no aparece en la cesta al ser entregada; el esclavo es acusado de haberlos comido, mostrándole como testimonio el papel con el mensaje escrito; el esclavo niega el hecho maldiciendo el tal papel "por ser un testigo falso y mentiroso". Tiempo después el mismo esclavo es enviado de nuevo con otra cesta de higos y con una carta en la que se informaba del número exacto de higos; el esclavo vuelve a comerse gran parte de los higos, "pero antes de tocarlos (para prevenir toda posible acusación) cogió la carta, y la escondió debajo de una gran piedra, tranquilizándose al pensar que si no lo veía comiéndose los higos, nunca podría referir nada de él; pero al ser ahora acusado con mayor fuerza que antes, confiesa su error, admirando la divinidad del papel, y para el futuro promete la mayor fidelidad en cada encargo." (Eco, 1992: 9). El esclavo había descubierto el uso que se hace de la escritura.

Sea o no ficción, esta narración le sirve a Eco para mostrar el poder de la representación escrita y los imaginarios que puede suscitar en quienes todavía no la reconocen como convención humana, pero que muy bien puede indicarnos ese halo de magia y de sorpresa en quien la descubre cuando accede a los dominios de la interpretación. Lo más relevante de dicha narración es sin duda la imagen de ser viviente que se le asigna a la escritura en el momento en que es objeto de lectura. Los textos hablan cuando tienen interlocutores, cuando hay alguien que coopera con ellos y los hace vivir; este alguien es esa figura

que llamamos lector; Umberto Eco habla de Lector Modelo, en esta perspectiva: el lector que se configura desde el texto.

Esta idea, la de los textos como hablantes, nos interesa desde el punto de vista pedagógico, pues sólo cuando los textos se asumen como instancias vivas, generadoras de provocaciones, puede decirse que hay lectura, lo cual nos hace interrogar por la manera como se abordan, y lo que ocurre con los textos en los ámbitos formales de la escuela: ¿hacen hablar? ¿propician el alegato? ¿hacen escribir?.

Comprender lo que es el acto de escribir implica preguntarse por su utilidad y por sus efectos de recepción en alguien que la interpreta. Ya se ha mostrado en muchas investigaciones (cfr, Ferreiro, Tolchinsky, Lerdner, Jolibert) cómo los niños aprenden a escribir con más gusto cuando saben que lo que escriben sirve para satisfacer a otros, para pedir algo que necesitan, para reclamar, para informar, para representar lo que sueñan, para mentir y manipular, para ponerle nombre a los dibujos y contar historias. Esto es contrario a la idea según la cual la escritura es hacer planas y copiar lo que está consignado en los libros.

El origen de la escritura constituye el mejor testimonio de las intenciones pragmáticas de su uso. De un lado, los grupos humanos han necesitado de un sistema de representación que permita retener en la perennidad los grandes eventos de la historia y las ocurrencias de los hombres y, de otro lado, han requerido de un sistema que garantice el cumplimiento de las normas jurídicas, que regule la compra y venta o intercambio de bienes, que respalde los acuerdos entre las personas y sirva para establecer un orden en las formas contables de la sociedad. Por eso, para los antiguos sumerios y egipcios la inscripción, en tabletas de arcilla, de figuras que representaban bultos o animales domésticos, que indicaban la cantidad de granos y de ganado, constituye una de las primeras formas de la búsqueda de dichos sistemas de representación: la escritura.

Puede decirse que la escritura, a través de su historia, ha estado determinada por dos extremos: como forma de poder y como forma de democratización; como forma de poder, porque las leyes están reguladas por la escritura de alguien o de un grupo y porque quien sabe escribir está en condiciones de manipular a otros; como forma de democratización, porque ejercer la escritura y, por supuesto, vivir la lectura, posibilita establecer relaciones intelectuales con alguna igualdad y, por lo tanto, estar en condiciones para la concertación y para reconocer las diferencias ideológicas con los demás; democratización porque la escritura y la lectura presuponen el acceso al saber universal y, con ello, la comprensión de los fenómenos sociales y políticos.

La escritura es un instrumento poderoso en la potenciación del pensamiento, por cuanto implica un razonamiento muy fino y un modo de obrar cauteloso y ponderado con las palabras. Es paradójico, sin embargo, que hoy cuando la educación escolarizada ha arrojado a índices amplios de la población, la escritura y la lectura no hayan logrado sus efectos en la formación de un ciudadano crítico y comprometido con un determinado proyecto social. Nos aterra, por ejemplo, que personas que han pasado por más de once años de escolaridad vendan por unos cuantos pesos el derecho al voto en las elecciones parlamentarias. Esto significa que se ha alfabetizado pero no se ha formado para leer y escribir y, en consecuencia, para actuar con criterio. Es decir, las personas han sido alfabetizadas pero no han logrado penetrar plenamente al mundo asombroso e interpelador de la lectura y la escritura.

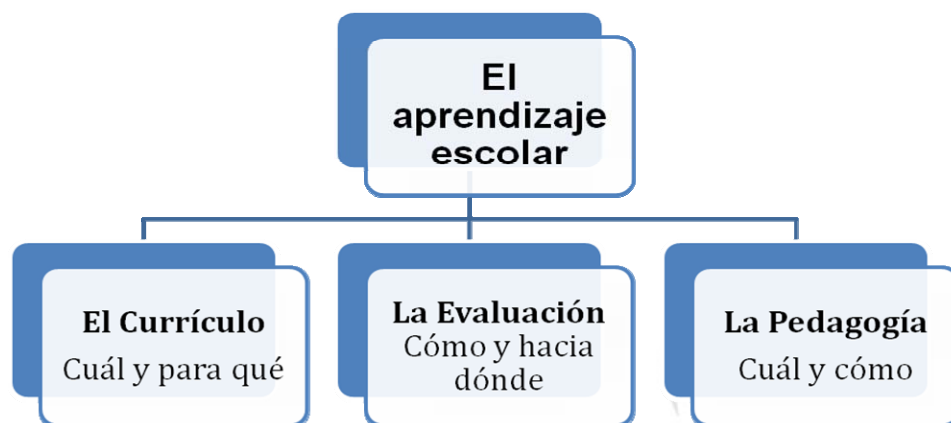
Así como Pedro Salinas afirma que lo que ha producido la escuela son leedores y no lectores, puede también decirse que ha producido escribidores y no escritores. Pero es un reto para quienes trabajamos en el espacio de la educación escolarizada (escuelas y universidades) propiciar el salto hacia la asunción

de la escritura, y con ésta la lectura, como una práctica decisiva para la emancipación y para dar sentido a la existencia. Con la lectura, y más con la escritura, se afianza el criterio y se gana en convicción y en argumentación. ¿Se orientan los proyectos de formación de los docentes hacia este propósito?

"La escritura es la manifestación íntegra del discurso", dice Ricoeur, pero además la escritura abre camino para comprender lo complejo que son los discursos mismos, por cuanto el leer, el leer como producción de sentido, es un acto de escritura, o un acto de representación simbólica en nuestro pensamiento: al leer reconstruimos la escritura que ha suscitado el acto del leer, porque el acto de interpretar es de cierto modo un acto de escritura.

2. SOBRE LA EVALUACIÓN DE LA ESCRITURA PARA EVALUAR LA LECTURA Y LA ESCRITURA

Del mismo modo como es necesario esclarecer lo que ha de entenderse por el rol del lector y el rol del escritor, es necesario definir lo que ha de entenderse por evaluación. La evaluación hace parte de la triada que constituye el dispositivo de la educación formal, si bien no se restringe estrictamente a esta triada:



Hay una relación de interdependencia entre el currículo, la evaluación y la pedagogía; la visión y las acciones de una, revierten en las otras: el enfoque sobre el currículo ha de enlazarse, con alguna consistencia teórica, con el enfoque sobre la evaluación y sobre la pedagogía. Sin embargo, es necesario distinguir entre el enfoque declarado (el que aparece, por ejemplo, en la legislación de cada país) y el enfoque realizado (el que aparece en las aulas). Es muy común hallar una consistencia en el enfoque declarado (cfr. SERCE-LLECE-UNESCO. 2004) mas no así en el enfoque que subyace en las acciones llevadas a cabo en las aulas, en las que prevalece un cierto eclecticismo o una desviación del enfoque declarado en la legislación. Por ejemplo: el enfoque declarado para la pedagogía en la mayoría de los países latinoamericanos pone su acento en el constructivismo para el aprendizaje de la lectura y la escritura, pero las acciones escolares recalcan en las actividades perceptivo-motrices, conductuales (con las planas) y la evaluación se restringe a dicha actividad (si la letra es bonita, si no emborrona, si no confunde las letras, si no se sale del renglón y mantiene una simetría...).

Con el objeto de develar las concepciones que regulan la actividad escolar es necesario ubicar a la evaluación en el centro de la triada, pues al analizar el modo de evaluar se puede inferir el enfoque sobre el currículo y sobre la pedagogía: dime cómo evalúas y te diré cómo concibes el currículo y la pedagogía. Definir la evaluación es un asunto complejo. La evaluación es una condición inherente al pensamiento y a la comunicación, al considerarla como un proceso valorativo (nadie asume una posición neutral respecto a lo que escucha o lee, pues de algún modo comprende desde valoraciones); por eso, la evaluación propicia interacciones; al contrario del examen canónico que pregunta por definiciones unilaterales (es la monología), la evaluación ha de concebirse en una dimensión polifónica, dialógica, en tanto que en el escenario de la interacción se representan los puntos contrapuestos sobre un determinado saber, las posibilidades y las conjeturas, las comparaciones y las analogías, todo ello posible gracias a las potencialidades de los sujetos.

Una característica relevante en esta mirada frente a la evaluación es el interés por devolver a una comunidad el análisis de lo que ella es, así como todo sujeto puede también darse cuenta de su mismidad cada vez que interactúa con otros (la valoración que cada uno hace). Dice Gardner que la evaluación “no se siente cómoda con el uso de instrumentos formales administrados en un entorno neutro y descontextualizado”, como lo son los instrumentos estandarizados, porque precisamente la evaluación polifónica se concentra en problemas puntuales, en aras de proponer soluciones y no en estigmatizar entre los excelentes, los buenos, los regulares y los deficientes. Sin embargo, es inevitable hoy el desarrollo de proyectos de evaluación externa para caracterizar los estados del aprendizaje alcanzado por los estudiantes y desde sus resultados definir planes para la cualificación de los procesos mismos. No se trata de la evaluación entonces para el ranking o para la estigmatización, sino de la evaluación para la interlocución con los maestros y con los directivos docentes.

Desde este ángulo de la evaluación la escritura aparece como una vía para evaluar los niveles de comprensión de lectura. Más que la estructura del test, necesaria para evaluaciones masivas, la producción escrita del evaluado pone al descubierto las propias lógicas del pensamiento en el modo de recuperar los conocimientos y lo aprendido desde los textos mismos; como lo hemos insinuado, la escritura, cuando funciona como tal, es el lugar donde convergen los textos que han sido transitados por quien escribe. La escritura pues permite identificar la apropiación semántica de los textos con los cuales se ha formado quien escribe y posibilita caracterizar los niveles de dominio en el modo de conducir un mensaje destinado a un posible lector. Estos dominios están relacionados con el uso adecuado de las categorías fundamentales de la lengua escrita, como lo es la concordancia gramatical, el uso pertinente de signos de puntuación y la cohesión entre las proposiciones, la discriminación y el enlace entre los párrafos...

Para la evaluación de la producción escrita, ya sea en el aula, o en los proyectos de evaluación externa, es necesario explicitar los criterios: qué se busca con la evaluación y cómo garantizar que sus resultados influyan en el proceso de transformación de los sujetos y de los proyectos educativos; el qué se busca, ha de conducir a la definición de plantillas, o rejillas, en las que se registran las categorías y los niveles de logro esperados. El cómo garantizar el impacto de los resultados tendría que orientarse hacia la formación de los docentes. En este punto es justo reconocer la importancia de hacer un balance sobre cómo los proyectos PISA y SERCE (Segundo Estudio Regional Comparativo Explicativo de la Calidad de la Educación, LLECE-UNESCO) le han apostado a la evaluación externa de la escritura. Hay aquí un camino

por investigar.

En el contexto escolar nos enfrentamos con el dilema sobre cómo evaluar la escritura de los estudiantes sin caer en una evaluación generalista o en una evaluación que no rebasa la impresión individual del maestro evaluador. Si hemos de aceptar que la subjetividad es inevitable en los procesos de valoración, esta subjetividad puede ser menor si se tienen unos criterios explícitos y una plantilla, que puede ser conocida incluso por los mismos evaluados. Frente a la masificación de las aulas, la plantilla (planilla le llaman en SERCE) constituye una posibilidad para que los estudiantes de grados superiores evalúen a los de los grados inferiores y con esta labor se evaluaría a la vez a los estudiantes evaluadores.

Toda evaluación académica que busca ser consistente y coherente con unos principios pedagógicos presupone una estructuración de las categorías (el objeto de la evaluación) y de los niveles de logro (los resultados esperados de la evaluación), apuntando hacia la caracterización de los estados de competencia en que se encuentran los sujetos evaluados; y este punto es fundamental: no se trata de evaluar para recriminar o para excluir sino para ayudarle al evaluado a identificar por dónde van sus procesos y a reflexionar sobre qué necesitaría para ponerse a tono con los demás, o simplemente para avanzar en sus propios procesos. Esta concepción invoca por la autonomía de los sujetos y por su capacidad para acceder a la metacognición: cómo estoy aprendiendo o por qué no logro aprender lo que me propongo aprender.

Una planilla, por ejemplo, como la de SERCE, orientada hacia la evaluación de la escritura para los niños de sexto grado, se inscribe en esta perspectiva. El grupo de lenguaje de SERCE, coordinado por Ana Atorresi, acordó evaluar la escritura considerando el paso del borrador al texto definitivo, según una consigna dada. La consigna solicitaba la elaboración de un texto descriptivo; este es otro punto central: el evaluador, ya sea el profesor de lenguaje, o de ciencias naturales, o de matemática o de otras áreas, ha de ser explícito en el tipo de texto que solicita sea producido por el estudiante: ¿texto narrativo? ¿texto descriptivo? ¿texto argumentativo? ¿texto explicativo? Observemos lo que se propuso evaluar, entre otros aspectos, SERCE:

Borrador								
1. ¿El niño escribió el borrador?	Si <input type="radio"/>	Si, pero borró lo escrito <input type="radio"/>	No <input type="radio"/>	3. ¿Siguió, en general, la consigna?	Si <input type="radio"/>	No <input type="radio"/>	N.C. <input type="radio"/>	
2. Si lo hizo, ¿lo redactó como un texto para sí?	Si <input type="radio"/>	No <input type="radio"/>	N.C. <input type="radio"/>	4. De las 4 ideas pedidas, ¿Cuántas planteó?	<input type="text"/>			N.C. <input type="radio"/>
Texto definitivo								
5. ¿El niño escribió el texto definitivo?	si <input type="radio"/>	Si, pero borró lo escrito <input type="radio"/>	No <input type="radio"/>	9. ¿El niño mantuvo los subtemas y la escritura planteados en el borrador?	si <input type="radio"/>	No <input type="radio"/>	N.C. <input type="radio"/>	
6. Si lo hizo, ¿mejoró la presentación respecto del borrador?	si <input type="radio"/>	No <input type="radio"/>	N.C. <input type="radio"/>	10. Independientemente de lo que haya hecho en el borrador, ¿se adecuó en general a la consigna?	si <input type="radio"/>	No <input type="radio"/>	N.C. <input type="radio"/>	
7. ¿El texto puede ser, en general, leído?	si <input type="radio"/>	No <input type="radio"/>	N.C. <input type="radio"/>	11. Si no escribió un texto acorde a la instrucción, ¿qué clase de texto escribió?				

8. Si puede ser leído, ¿excluye los problemas de desciframiento?	si	No	N.C	<input type="radio"/> Una copia completa o parcial, textual o parafraseada, de la instrucción	Una anécdota	<input type="radio"/>
	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/> Un cuento	Otra descripción	<input type="radio"/>
					Otro	<input type="radio"/>
					Un no-texto	<input type="radio"/>
					N.C.	<input type="radio"/>

Una señal pedagógica de la estrategia de evaluación de SERCE está en el reconocimiento de la importancia del borrador (es concebir la escritura como proceso) en el acto de escribir, así como el carácter cualitativo de la evaluación (se trata de identificar el nivel de competencia en escritura) y su horizonte en el aprendizaje: el niño reconocerá lo logrado y lo que está pendiente por aprender. Se espera que desde una evaluación como esta el profesor o la profesora contribuyan a que el mismo estudiante sea quien reconozca el estado de su saber.

Por otro lado, como se trata de contextos académicos, y como una manera de controlar los límites de los objetos evaluados, es deseable evaluar la producción escrita considerando el uso de textos-fuentes, como lo mostraremos en el ejemplo que sigue. Evaluaciones como estas –de la lectura hacia la escritura y viceversa- son necesarias para caracterizar los niveles de dominio de los textos leídos y los niveles de uso de la lengua escrita. Señalamos esto para reducir las tensiones cuando el docente imagina que tiene que evaluar formalmente todo lo que el estudiante hace. Desde nuestro punto de vista las acciones humanas son objeto de valoración continua pero no necesariamente de evaluaciones puntuales y controladas.

3. ESTUDIO DE UN CASO

Entre los años 1998 a 2002, en Bogotá se desarrolló el proyecto “Evaluación Censal de Competencias Básicas”. El proyecto fue liderado por el Grupo de Investigación en Evaluación, de la Universidad Nacional de Colombia; el proyecto se propuso:

1. Elaborar pruebas en las que se integraran áreas básicas, como lenguaje y matemáticas, para tercero y quinto grado, y lenguaje, matemáticas y ciencias, para séptimo y noveno grado, teniendo como referente conceptual las competencias insinuadas en los lineamientos curriculares del Ministerio de Educación. Dichas pruebas se diseñaron con una parte cerrada (estructura del test) y una parte abierta (producción de un texto). Para tercero y quinto grado se tomó como referente de la prueba una narrativa icónica (un dibujo en el que se representan eventos comunicativos); para séptimo y noveno, se toma como referente un periódico, en el que aparecen secciones sobre ciencia y tecnología, salud y nutrición, pasatiempos, deportes, arte y literatura.
2. Asumir la escritura como una práctica inherente a los distintos campos del conocimiento (matemática, ciencias, lenguaje), identificando rejillas específicas para el análisis de la escritura en cada una de las áreas; ello por cuanto se buscaba saber cómo la escritura de los estudiantes mostraba haber alcanzado los aprendizajes de aspectos relevantes en dichas áreas, pero también para mostrar cómo con un mismo escrito podemos evaluar diversos campos de conocimiento.
3. Hacer públicos los instrumentos luego de su aplicación, con el objeto de usarlos como material pedagógico en el aula (que los estudiantes trabajen de nuevo las pruebas con sus docentes) y

que sirvan para analizar con los docentes por qué se evalúa lo que se evalúa e inclusive puedan identificar errores en las pruebas.

4. Desarrollar procesos de acompañamiento in situ a las instituciones, orientando la formación permanente de los docentes, propiciando condiciones para la investigación cooperativa y buscando otras formas de evaluación, distintas al test, desde las cuales se puedan caracterizar otras competencias (la comunicación oral, la escucha, la expresión gestual...).
5. Promover la investigación entre los maestros, contrastando los diseños y los resultados de la evaluación interna (la de la cotidianidad en el aula) con la evaluación externa (la que es diseñada por agentes externos de la escuela): ¿coinciden los resultados?

Los estudiantes recibieron un cuadernillo para responder (marcando una respuesta) según una serie de textos contenidos en un periódico escolar (toda la prueba gira alrededor del periódico). Al final del cuadernillo aparece un espacio en blanco con el encabezamiento de la consigna: "Teniendo en cuenta lo que has leído en el periódico Lo que Somos nos gustaría saber lo que tú piensas sobre las relaciones entre el suelo, la producción de alimentos y lo que puede esperar la creciente población humana."

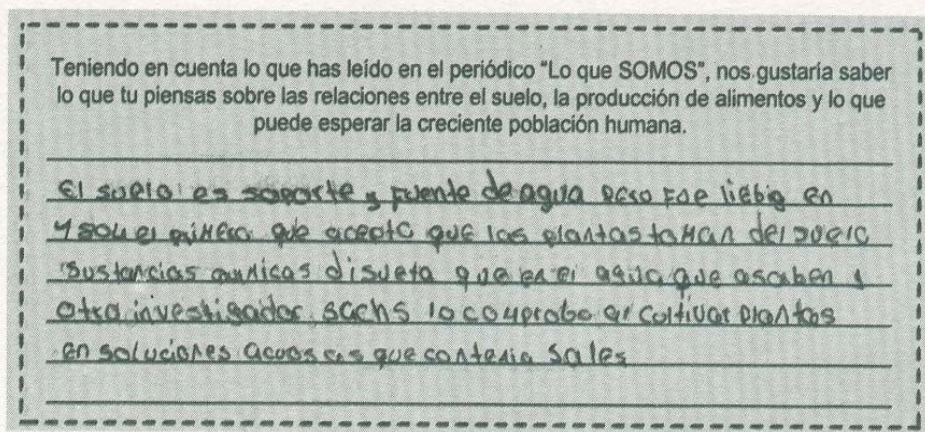
A continuación presentamos la rejilla desde la cual se evaluaron los escritos de los estudiantes del grado séptimo (segundo año de secundaria), desde la perspectiva de las ciencias naturales y mostramos los ejemplos correspondientes.

Categoría	Condiciones	Nivel			
Referencia a la temática planteada.	Hacer referencia a una de las variables: el suelo, la producción de alimentos o la creciente población humana.	A			
Relación entre dos variables.	Establecer una relación pertinente entre dos de las variables.		B		
Relación entre tres variables.	Establecer una relación pertinente entre las tres variables.			C	
Relación entre tres variables y fundamentación en texto base.	Establecer de manera explícita la relación pertinente entre las tres variables y fundamentar en las ideas planteadas en los textos de base.				D

En la rejilla se identifican las categorías (son los objetos que se evalúan), las condiciones (son los referentes que muestran el cumplimiento con la categoría) y el nivel (es la identificación del logro del

estudiante), teniendo en cuenta el mínimo (A) y lo máximo (D) en la valoración. Este último nivel incluye a todos los anteriores. Veamos los ejemplos y sus análisis.

Texto I



El texto del estudiante dice:

"El suelo es soporte, fuente de agua pero fue Liebig en 1804 el primero que aceptó que las plantas toman del suelo sustancias químicas disueltas que en el agua que absorben y otro investigador Sachs lo comprobó al cultivar plantas en soluciones acuosas que contenían sales"

Un fragmento del texto que aparece en el periódico dice:

"Se sabía de tiempo atrás, que el suelo es soporte y fuente de agua, pero fue Liebig, en 1804, el primero que aceptó que las plantas toman del suelo sustancias químicas disueltas en el agua que absorben y otro investigador, Sachs, lo comprobó al cultivar plantas en soluciones acuosas que contenían sales que les suministraban potasio, calcio, fósforo, azufre, hierro, magnesio y nitrógeno."

El texto del estudiante es pues una transcripción afanada del fragmento del periódico. En estos casos no es posible evaluar la lectura del texto periodístico a partir del escrito, por cuanto no hay producción del estudiante, de lo cual se concluye que al no haber lectura del texto es imposible que haya escritura; la transcripción mecánica no se constituye en escritura. El caso ha de servir para ejemplificar, en el contexto del aula, la tendencia de los estudiantes a tomar información de Internet o copiar de las enciclopedias para presentarla como propia. Desde el ámbito pedagógico el caso ha de utilizarse para mostrar la imposibilidad de hacer la evaluación al estudiante por la ausencia de producción: no podemos decir nada de cómo escribe el estudiante, más allá de caracterizar sus trazos y sus grafías, pero esto, como lo hemos señalado, no es lo fundamental. De acuerdo con la rejilla el escrito no se acopla a ninguno de los niveles.

Texto 2

Teniendo en cuenta lo que has leído en el periódico "Lo que SOMOS", nos gustaría saber lo que tu piensas sobre las relaciones entre el suelo, la producción de alimentos y lo que puede esperar la creciente población humana.

que el suelo es un recurso
que nosotros debemos cuidar
porque sin el no podemos
vivir y que en el suelo
podemos sembrar lo que
queramos y que los
alimentos los comamos
porque nos dan energía
proteínas y calcio y si
no comamos los alimentos
podemos sufrir enfermedades
y que sepamos aprovechar
lo que tenemos en este
mundo

El texto del estudiante dice:

"que el suelo es un recurso que nosotros debemos cuidar porque sin el no podemos vivir y que en el suelo podemos sembrar lo que queremos y que los alimentos los comamos porque nos dan energía proteínas y calcio y si no comamos los alimentos podemos sufrir enfermedades y que sepamos aprovechar lo que tenemos en este mundo".

El texto se inscribiría en el nivel A, pues "hace referencia a la temática planteada" y es en sí una producción escrita del estudiante (se representan unas ideas). En el escrito aparecen dos variables (el suelo y los alimentos) pero sin un enlace lógico-semántico. El periódico presentaba además un artículo sobre la composición de los alimentos que el hombre consume; el estudiante condensa significados de dos textos y plantea una idea general sobre el suelo y sobre los alimentos. Desde el punto de vista lingüístico se observa la presencia de registros orales (y.. y... y...) y la ausencia de signos de puntuación. El escrito recoge aspectos fragmentados de lo comprendido en la lectura del periódico.

Texto 3

Teniendo en cuenta lo que has leído en el periódico "Lo que SOMOS", nos gustaría saber lo que tu piensas sobre las relaciones entre el suelo, la producción de alimentos y lo que puede esperar la creciente población humana.

El suelo y la producción de alimentos se relacionan porque el suelo es la base para que se desarrollen los alimentos y la población se alimenta de este pero el suelo se va gastando y se va perdiendo la fertilidad del suelo y las plantas y vegetales se van acabando y teniendo un crecimiento pobre y lento ya que todos nosotros no sabemos tratar los suelos podemos hacer forestación, rotación de cultivos, drenajes en fin con tal de saber tratar los suelos y manejarlos.

El texto del estudiante dice:

"El suelo y la producción de alimentos se relacionan porque el suelo es la base para que se desarrollen los alimentos y la población se alimenta de este pero el suelo se va gastando y se va perdiendo la fertilidad del suelo y las plantas y vegetales se van acabando y teniendo un crecimiento pobre y lento ya que todos nosotros no sabemos tratar los suelos podemos hacer forestación, rotación de cultivos, drenajes en fin con tal de saber tratar los suelos y manejarlos".

De acuerdo con la rejilla, en el escrito se logra establecer la relación pertinente entre dos variables (el suelo y los alimentos); a diferencia del escrito anterior en el que se introduce la variable del suelo pero no se la relaciona explícitamente con los alimentos, aquí se introduce al comienzo el tema del suelo, se establece inmediatamente la relación con los alimentos y desemboca después en la conclusión, según la cual como no sabemos tratar los suelos, es necesario hacer reforestación, rotar los cultivos, hacer drenajes para tratar mejor los suelos. Aparecen registros semánticos de dos artículos del periódico; es decir, revela un cierto nivel de lectura de los artículos del periódico. Desde el punto de vista lingüístico prevalece la repetición de la conjunción "y", como una marca de la oralidad, y solo aparece el punto final como signo de puntuación. El escrito se inscribiría en el nivel B.

Texto 4

Teniendo en cuenta lo que has leído en el periódico "Lo que SOMOS", nos gustaría saber lo que tú piensas sobre las relaciones entre el suelo, la producción de alimentos y lo que puede esperar la creciente población humana.

Yo pienso que el suelo es parte importante para nuestro desarrollo económico ya que a través de ella podemos adquirir los alimentos necesarios para nuestra vida diaria; debemos cuidar el suelo y con un buen cuidado y manejo será muy fértil.

El planeta tiene sobrepoblación. La gente debería de caer en cuenta y controlar la sobrepoblación; en Japón y China se presenta la sobrepoblación mas grande y si seguimos así acabaremos sin espacio y mal.

El escrito del estudiante dice:

"Yo pienso que el suelo es parte importante para nuestro desarrollo económico ya que a través de él podemos adquirir los alimentos necesarios para nuestra vida diaria; debemos cuidar el suelo y con un buen cuidado y manejo será muy fértil.

El planeta tiene sobrepoblación. La gente debería de caer en cuenta y controlar la sobrepoblación; en Japón y China se presenta la sobrepoblación mas grande y si seguimos así acabaremos sin espacio y mal".

El estudiante introduce las tres variables estableciendo relaciones: el suelo, los alimentos producidos y la sobrepoblación. Construye dos párrafos, ubicando en el primero la relación entre el suelo y los alimentos y, en el segundo, referencia la sobrepoblación introduciendo un saber que no aparece en los textos del periódico: Japón y China. Desde el punto de vista lingüístico hay un uso adecuado de los signos de puntuación: el punto y coma, el punto aparte y el punto final. Como se ha podido observar en otros estudios sobre la escritura de los estudiantes, se constata cómo la consistencia de la escritura está determinada en gran parte por el uso pertinente de los signos de puntuación. El uso adecuado de los signos de puntuación permite observar un dominio de la competencia textual: saber conducir el mensaje desde una coherencia lineal y global. El texto cumple con las condiciones del nivel C.

Texto 5

Teniendo en cuenta lo que has leído en el periódico "Lo que SOMOS", nos gustaría saber lo que tu piensas sobre las relaciones entre el suelo, la producción de alimentos y lo que puede esperar la creciente población humana.

La verdad es muy preocupante lo que se espera para la población humana puesto que nos estamos autodestruyendo al realizar la agricultura a una gran escala, ocasionando así erosión y la casi destrucción del suelo, generando menos cultivos de plantas y por lo tanto menor producción de alimentos y sus variedades. Es muy posible que nuestro planeta en sus partes más bellas quede sólo desierto, y el problema no llega allí, también afecta el ecosistema natural permitiendo la extinción de especies. Además, si no hay alimento el ser hombre tendrá que recurrir a sólo carnes y lácteos, desequilibrando un consumo normal y el desarrollo y crecimiento normal. Ahora, esto también afecta la economía ya que las frutas y vegetales subirán rápidamente su costo, al ser estos productos de escasos.

El texto del estudiante dice:

"La verdad, es muy preocupante lo que se espera para la población humana puesto que nos estamos destruyendo al realizar la agricultura a una gran escala, ocasionando así erosión y la casi destrucción del suelo, generando menos cultivos de plantas y por lo tanto menor producción de alimentos y sus variedades. Es muy posible que nuestro planeta en sus partes más bellas quede sólo desierto, y el problema no llega allí, también afecta al ecosistema natural permitiendo la extinción de especies. Además, si no hay alimento el ser hombre tendrá que recurrir a sólo carnes y lácteos, desequilibrando un consumo normal y el desarrollo y crecimiento normal. Ahora, esto también afecta la economía ya que las frutas y vegetales subirán rápidamente su costo, al ser estos productos de escasos."

En un apartado del artículo "El rincón agrícola", que aparece en el periódico, se observa el enunciado: "Una consecuencia de la agricultura a gran escala ha sido la erosión de los suelos". En el escrito del estudiante hallamos: "... nos estamos destruyendo al realizar la agricultura a una gran escala, ocasionando así erosión y la casi destrucción del suelo". El estudiante realiza una paráfrasis del texto que ha leído —es lo más recurrente en los escritores expertos: la paráfrasis o la glosa— y además hace un uso adecuado de la inferencia (la inevitable destrucción del suelo, por efecto de la agricultura a gran escala). El estudiante introduce las tres variables y se apoya directamente en el texto al citar un apartado de uno de los textos que aparece en el periódico; no es necesario el uso de las comillas por cuanto parafrasea dicho apartado.

Desde el punto de vista lingüístico-textual el escrito se caracteriza por la introducción de conectores que posibilitan un estilo genuino de la escritura y una presencia de la argumentación: "puesto que", "es muy posible que", "además", "ahora". Introduce juicios valorativos sobre el fenómeno: "nuestro planeta en sus partes más bellas". Hay asimismo un uso adecuado de los signos de puntuación: recurre a la coma y al punto. El escrito cumple con las condiciones para el nivel D, de acuerdo con la rejilla.

4. A MANERA DE CONCLUSIÓN

El horizonte pedagógico desde la metacomunicación presupone tener la disposición del profesor para leer estructuras profundas en los escritos de sus estudiantes y no anclar la mirada únicamente en las estructuras superficiales (la deficiente redacción, se dice). Leer estructuras profundas en discursos escritos con anomalías sintácticas revela un compromiso y una actitud pedagógica consecuente con el oficio de ser maestro: si hay estructura profunda, esto es, si hay construcción de sentido, hallamos el punto de partida para la interacción con el estudiante; en este reconocimiento de la estructura profunda el profesor asume el rol de un corrector de estilo porque aspira a que sus estudiantes descubran la anomalía y procedan con la reescritura, algo inherente a todos los que escribimos: a todos nos corrigen. Pero de nuevo aparece el dilema sobre qué tanto escriben los profesores, para sus clases como para los medios académicos: revistas, portales y páginas web y libros. Hemos de considerar que sólo quien escribe con una cierta regularidad puede fungir de interlocutor con otros que también escriben. El compromiso de ayudar a cualificar, progresivamente, la escritura de los estudiantes compromete a los profesores de todas las áreas curriculares porque se aprende a leer y a escribir según los códigos recurrentes en cada disciplina (en cada área curricular) y no como se cree le competiría únicamente al área de lengua.

Dice Gardner que "en lugar de imponerse de forma externa en momentos inoportunos durante el curso, la evaluación tendría que formar parte del entorno natural del aprendizaje." (1998: 188). En "el entorno natural del aprendizaje" los saberes dialogan y se representan en los esquemas, o mapas conceptuales. La fuerza de la escritura está orientada siempre por unos esquemas previos y estos esquemas devienen de la experiencia de la lectura: según como uno lee, uno escribe, si bien el asunto no es tan taxativo. En esto radica la potencia de la pedagogía: se trata de analizar cómo funcionan los textos para aprender a dominarlos y a producirlos; no se trata de enseñar a leer y a escribir, porque esto no es posible; se trata de aprender a interpretar los múltiples textos de la cultura en un proceso de interlocución para luego producir otros textos; un interlocutor fundamental, insistamos, es el maestro.

Finalmente, es necesario reiterar lo inútil de ofrecer cursos compensatorios para resolver los problemas con la escritura; estos problemas no se resuelven con cursos remediales sino viviendo la escritura desde el interés por un campo y desde la interacción en el seno de una comunidad —una comunidad es la que se constituye en un salón de clase, por ejemplo—. Quien escribe y no tiene interlocutores —y el maestro debe serlo— difícilmente podrá avanzar y cualificar los modos de leer y los modos de escribir. La escritura de alguien deja ver los modos como lee, porque el texto producido revela las voces desde donde habla, lo cual quiere decir que el texto no es puro: está impregnado de voces.

La evaluación tendría que acometer (analizar) la fuerza polifónica de la escritura (cfr Bajtin, 1982) y esto compromete al profesor como experto en lectura —es lo que se supone— quien podrá rastrear esas voces-fuentes en los escritos de sus estudiantes y valorarlos según son entretejidas en la escritura. Es el profesor, como experto lector, quien puede sugerir los textos-fuentes para desde ellos producir el nuevo texto: el del estudiante, pero también el suyo, como profesor: esto es hacer comunidad académica. No hay otro modo para contribuir en la cualificación progresiva de la escritura de los estudiantes si no es desde la interlocución y en una relación de pares académicos, más aún hoy cuando los estudiantes parecen ir más lejos que sus propios profesores por el contacto cotidiano que tienen con el computador.

El potencial intuitivo, inherente a la formación pedagógica del profesor, es definitivo para saber cuándo el estudiante copia, bajando información de Internet y sin procesarla; en estos casos el problema no es del estudiante, que lo hace desde un razonamiento particular, sino del profesor que no lo detiene para explicarle que como tal escritura no es suya no es posible evaluarlo. Es grave que los profesores dejen pasar como válidos los escritos bajados de Internet y ser presentados por los estudiantes como suyos. Cuando esto ocurre se revela cómo el profesor no es un experto ni en lectura ni en escritura; seguramente se ha equivocado de profesión, porque esta es una condición de todo profesor, sea de física, de química, de matemática, de historia o de filosofía: es idóneo en el acceso a los textos y en el modo de producirlos.

Para asumir la evaluación de la lectura y la escritura desde esta perspectiva se requiere entonces que los maestros vivan y sientan la necesidad de escribir, para lo cual es una condición participar en proyectos, tener medios disponibles para publicar (revistas, periódicos, portales virtuales, edición de libros) y socializar a través de ponencias en congresos y simposios las reflexiones sobre las prácticas o sobre las disciplinas del currículo. De este modo los estudiantes también sentirán esa necesidad y asumirán la escritura no como un deber sino como un deseo. También ellos participarán en simposios y congresos con ponencias propias y fundarán revistas y periódicos, así como páginas web: es el horizonte pragmático de la escritura como representación de la lectura.

REFERENCIAS BIBLIOGRÁFICAS

- Atorresi, A. et. al. (2009). *Aportes a la enseñanza de la lectura*. Santiago de Chile: UNESCO
- Bajtin, M. (1982). *Estética de la creación verbal*. México: Siglo XXI.
- Bajtin, M. (1997). *Hacia una filosofía del acto ético*. Barcelona: Anthopos.
- Calsamiglia, H. y Tusón, A. (1999). *Las cosas del decir*. Barcelona: Ariel
- Cassany, D. (2007). *Afilas el lapicero*. Barcelona: Anagrama.
- Ducrot, O. (1988). *Polifonía y argumentación*, Cali: Univalle.
- Eco, U. (1981). *Lector in fabula*. Barcelona: Lumen.
- Eco, U. (1988). *De los espejos y otros ensayos*. Barcelona: Lumen.
- Eco, U. (1992). *Los límites de la interpretación*. Barcelona: Lumen.
- Ferreiro, E. (1989). *Los hijos del analfabetismo*. México: Siglo XXI
- Freire, P. (1997). *La educación en la ciudad*. México: Siglo XXI

- Garrido, F. (1999). *El buen lector se hace, no nace*. México: Planeta Ariel
- Gardner, H. (2000). *La educación de la mente y el conocimiento de las disciplinas*. Barcelona: Paidós
- Gardner, H. (2001). *La inteligencia reformulada*. Barcelona: Paidós
- Jurado, F. (2001). Formación del profesorado y prácticas educativas en el área del lenguaje y de la literatura en América Latina". *Revista Textos*, 27.
- Jurado, F. (1997) *Investigación, escritura y educación*. Bogotá: Universidad Nacional.
- Jurado, F. y Sánchez, V. (2001). La práctica como motor de la transformación de la formación docente. En: *La formación docente en América Latina*. Bogotá: Universidad Nacional de Colombia – Cooperativa Editorial Magisterio – Red Latinoamericana
- Jolibert, J. (1991). *Formar niños productores de textos*. Santiago de Chile: Dolmen.
- Lerdner, D. (2001). *Leer y escribir en la escuela: lo real, lo posible y lo necesario*. México: FCE.
- Lomas, C. (comp.) (2001). *¿Educar o segregar? Materiales para la transformación de la educación secundaria*. Bogotá: Universidad Nacional de Colombia – Cooperativa Editorial Magisterio
- Marina, J.A. (1993). *Teoría de la inteligencia creadora*. Barcelona: Anagrama.
- Ricoeur, P. (1995). *Teoría de la interpretación*. México: Siglo XXI-Universidad Iberoamericana.
- Ricoeur, P. (1985). *Tiempo y narración* (t. I), trad. Agustín Neira, México: Siglo XXI.
- Ricoeur, P. (1985). *Tiempo y narración* (t. III), trad. Agustín Neira, México: Siglo XXI.
- Secretaría de Educación de Bogotá (1999). *Resultados. Evaluación de competencias básicas en lenguaje, matemática y ciencias*. Bogotá: SED.
- SERCE-LLECE-UNESCO (2004). *Análisis curricular*. Santiago de Chile: UNESCO.
- Van-Dijk, T. (1982). *La ciencia del texto*. Barcelona: Paidós.
- Tolchinsky, L. (1993). *Aprendizaje del lenguaje escrito*. Barcelona: Anthopos.
- Voloshinov, V. (1992). *El marxismo y la filosofía del lenguaje*. Madrid: Alianza.



ISSN: 1989-0397

LA EVALUACIÓN DE APRENDIZAJES DE PERSONAS JÓVENES Y ADULTAS

María Eugenia Letelier

Revista Iberoamericana de Evaluación Educativa 2009 - Volumen 2, Número 1

<http://www.rinace.net/riee/numeros/vol2-num1/art8.pdf>

Fecha de recepción: 15 de marzo de 2009

Fecha de comunicación de dictamen: 18 de abril de 2009

Fecha de aceptación: 18 de abril de 2009

1. AVANCES EN LA EVALUACIÓN DE APRENDIZAJES DE PERSONAS JÓVENES Y ADULTAS: MEDICIONES Y EXPERIENCIAS INTERNACIONALES

A diferencia de la educación de niños y jóvenes, en que la mayoría de los países cuentan con sistemas de medición de la calidad y existen experiencias de participación periódicas en pruebas internacionales como son PISA, TIMSS y PIRLS, en educación de personas jóvenes y adultas son pocas las instancias de generación de conocimiento a través de intercambio y de participación en este tipo de pruebas.

En el ámbito internacional, las mediciones que consideran a la población adulta son escasas, puntuales y se han implementado en un número restringido de países. Entre ellas destacan el IALS (*International Adult Literacy Survey*), proyecto impulsado por la Organización para la Cooperación y Desarrollo Económico (OCDE)¹ y el PIAAC (*Programme for the International Assessment of Adult Competencia*), reciente estudio que está en una fase inicial de desarrollo, orientado a evaluar el nivel y la distribución de las competencias de los adultos; en ambos casos, Chile ha sido el único país latinoamericano participante.

Entre 1996 y 1998, la Oficina Regional de la UNESCO (OREALC) impulsó una investigación para medir competencias de los adultos, que fue aplicada en ciudades de siete países de América Latina. Con una metodología similar al IALS, pero adaptando los instrumentos al contexto latinoamericano, se aplicaron un conjunto de ítems, que “permitieron identificar niveles diferentes de la población en el desarrollo de las competencias demostradas en la resolución de tareas en formatos de prosa, documentos o esquemas y matemática”. (Infante: 2000)

Otra investigación que se encuentran en una fase inicial es el proyecto *Programa de Evaluación y Monitoreo de la Alfabetización (LAMP*, por su sigla en inglés) impulsado por el Instituto de Estadísticas de la UNESCO que también se propone medir las competencias de los adultos, incluyendo un número pequeño de países, entre ellos El Salvador y Paraguay como únicos países de habla castellana.

Todas estas investigaciones han buscado medir las competencias de los adultos de manera directa, comprobando las tareas que las personas son capaces de desarrollar en cuanto a comprensión lectora, cálculo y manejo de información presentada en diferentes formatos. Todas se aplican a una muestra de hogares, generalmente, a personas entre 15 y 65 años, por tanto, no tienen como propósito medir ni comparar la calidad de los aprendizajes logrados por los alumnos y alumnas que participan en alguna de las modalidades de educación de adultos.

Estas investigaciones, a las que se suman algunas experiencias nacionales como es el caso del Indicador de Alfabetismo Funcional –INAF- desarrollado en Sao Paulo, (Brasil) han contribuido a actualizar concepciones en cuanto a la medición del alfabetismo de la población y hacer comparables los resultados entre los países que participan. Pero, son insuficientes para apoyar la reflexión pedagógica y la construcción de sistemas de evaluación de aprendizajes específicos para la educación de adultos.

¹ Entre 1994 y 1996 se desarrolló la Primera Encuesta Internacional realizada en 12 países miembros de la OECD proyecto coordinado por Statistics (Canadá) y Education Testing Service (Estados Unidos). En 1998 el Proyecto se amplió a otros 10 países, entre los cuales Chile fue el único país de lengua castellana

La misma conclusión se puede extraer cuando se buscan referencias en programas no-formales de adultos y, especialmente, de alfabetización. Aquellos programas que cuentan con mediciones relacionadas con el impacto en los aprendizajes de los participantes, generalmente, están ligados a los requerimientos de la cooperación internacional, como es el caso de los PAEBA's (Programa de Educación Básica de Jóvenes y Adultos, financiados por la cooperación española), más recientemente, se pueden encontrar mediciones nacionales como es el Programa Brasil Alfabetizado y del Proyecto SESI (Serviço Social da Indústria), también en Brasil. En Chile, además, se ha instalado un procedimiento externo de evaluación de los aprendizajes de las personas que participan de la Campaña de Alfabetización Contigo Aprendo que, en caso de aprobar, conduce a la certificación del primer nivel de Educación Básica.

Aunque con rezago, en casi todos los países de América Latina se han iniciado procesos de reforma de educación de adultos, muchas de estas reformas han impulsado programas flexibles de nivelación de estudios que buscan adecuarse de mejor manera a las condiciones de vida y de trabajo de las personas adultas. Para que sea posible la flexibilización de la oferta y, al mismo tiempo se cautele la calidad de los aprendizajes que se certifican, algunos países han construido mecanismos de evaluación externo al proceso educativo.

Con este propósito, en algunos países de América Latina, se han creado sistemas de evaluación de aprendizajes y certificación de estudios. Entre ellos, destacan el Sistema Automatizado de Seguimiento y Acreditación (SASA) de México, sistema a través del cual se elaboran y aplican exámenes de diagnóstico y final a las personas que se inscriben en el Modelo de Educación Para la Vida y el Trabajo (MEVyT), dependiente del Instituto Nacional de Educación de Adultos (INEA) y el Sistema Nacional de Evaluación de Aprendizajes y Certificación de Estudios de Chile, que evalúa a los alumnos y alumnas integrados en programas de nivelación flexible que buscan certificar sus estudios de Educación Básica o Media. Por último, en un nivel más incipiente, en El Salvador, también se construye una experiencia de evaluación externa y orientada a la certificación de estudios para jóvenes y adultos que participan del programa EDUCAME.

A nivel internacional, el tema ha estado presente en las reuniones preparatorias a la Conferencia Internacional de Alfabetización y Educación de Adultos (CONFINTEA) que en su sexta versión se desarrollará en mayo de 2009. En el documento final de recomendaciones de la reunión regional de países de América Latina y El Caribe, se propone « desarrollar un sistema de evaluación, reporte, registro y monitoreo con parámetros internacionales que posibiliten la formulación de políticas a partir de la evaluación de los procesos, sistemas y métodos, que asegure la certificación, acreditación y homologación de los conocimientos y habilidades » (UIL/UNESCO 2008).

La preocupación por orientar políticas de evaluación de adultos entre los países que forman parte del Programa Iberoamericano de Alfabetización y Educación Básica de Adultos (PIA) impulsó a la Organización de Estados Iberoamericanos (OEI) en conjunto con el Gobierno de Chile, a realizar un seminario internacional con el propósito de "promover el debate académico y el intercambio de experiencias sobre construcción de sistemas de evaluación de aprendizajes de personas jóvenes y adultas,

en el marco de las políticas de educación de adultos”², sus contenidos darán origen a una publicación que contribuirá a alcanzar los objetivos propuestos.

Estas iniciativas permiten tener un panorama general acerca del interés que está adquiriendo la evaluación de los aprendizajes en Educación de Adultos. Se puede afirmar que este es un campo en evolución, que por su complejidad, requiere de reflexión académica y apoyo técnico que permita implementar procedimientos adecuados a los propósitos de la evaluación y a las características de los programas y de las personas a las cuales se dirigen.

2. ¿QUÉ ASPECTOS ESPECÍFICOS SE DEBEN CONSIDERAR EN LA EVALUACIÓN DE APRENDIZAJES DE PERSONAS JÓVENES Y ADULTAS?

Existen varios motivos para diferenciar la evaluación de los aprendizajes de las personas jóvenes y adultas, respecto a la que se realiza con niños y jóvenes del sistema educativo regular. En primer lugar, la diversidad y mayor flexibilidad de la oferta educativa dirigida a las personas adultas; en segundo lugar, quienes asisten a la educación de personas jóvenes y adultas tienen características que los diferencian de los niños no sólo en cuanto a ciclo y condiciones de vida sino, especialmente, en cuanto a las estrategias de aprendizaje y motivaciones que tienen para aprender; en tercer lugar, aplicar a los adultos instrumentos de evaluación diseñados y elaborados para medir los aprendizajes de los niños limita el desarrollo cognitivo y afecta la autoestima de las personas.

2.1. La diversidad de los programas y el propósito de la evaluación

Sin duda que un rasgo propio de la educación de adultos es la diversidad de la oferta educativa, diversidad que refleja la propia heterogeneidad de las poblaciones que participan en los procesos educativos. Esta diversidad incluye: a) programas formales que tradicionalmente son parte de la oferta escolar regular, la que con frecuencia incorpora a jóvenes desertores; b) programas no formales que conducen a certificación de estudios, pero cuya característica fundamental es la flexibilidad de la oferta educativa, razón por la cual se desarrollan en una diversidad de poblaciones y contextos socio-culturales; c) programas informales, que suelen ser de corta duración y que tienen objetivos ligados a procesos de capacitación y/o desarrollo sobre algún tema significativo para la vida de las personas adultas. Junto a esta diversidad de ofertas, están las campañas y programas de alfabetización que juegan un papel fundamental no sólo como acciones educativas, sino como acciones sociales que suelen movilizar una gran cantidad de recursos humanos y materiales.

En todas estas modalidades y programas educativos es necesario evaluar el aprendizaje de las personas. En los programas formales o informales que otorgan certificación de estudios, las evaluaciones deberán cumplir requisitos asociados a la normativa educativa y ser consistentes con los objetivos y contenidos del currículum que acredita. Si el proceso educativo se desarrolla en el ámbito escolar, será el docente en el aula quien evalúe con fines diagnóstico, formativo o sumativo; en cambio, en los programas no formales

² El seminario, realizado en diciembre del año 2008, contó con la intervención de seis prestigiosos expertos que dictaron conferencias vinculadas a la evaluación de aprendizajes y la construcción de sistemas de evaluación en Educación de Adultos, y la participación de delegados de catorce países iberoamericanos, algunos de los cuales dieron a conocer experiencias que se desarrollan en sus respectivos países.

que conducen a acreditación escolar, la evaluación suele medir resultados finales de aprendizajes y los instrumentos son aplicados con un procedimiento externo al proceso educativo.

Los programas que no conducen a acreditación escolar, muestran una amplia gama de modalidades y temáticas, están aquellos relacionados con el mundo de la capacitación laboral, que en algunos países ha llevado a impulsar complejos sistemas de reconocimiento de estos aprendizajes. También se pueden incluir una variedad de programas organizados en temas como salud, cuidado del medioambiente, ciudadanía, derechos de los consumidores, etc. Todos ellos conforman un conjunto de iniciativas que se consideran parte de la educación de adultos, desde el punto de vista de la evaluación estas acciones requieren de seguimiento y valoración de los aprendizajes que las personas han alcanzado.

De acuerdo a la modalidad, propósito e impacto que tienen estos programas, será necesario dar respuesta a algunas preguntas básicas relacionadas con todo proceso de evaluación, entre ellas: qué se evaluará, con qué tipo de instrumentos, qué tipo y cantidad de ítems se aplicarán, cómo y en qué momento se aplicará, cómo se informan los resultados a los participantes.

2.2. Las características de la población

Toda evaluación debe incorporar una visión de la población que se evalúa. Por ello, en los procedimientos de evaluación de adultos, será necesario incorporar aspectos propios de los jóvenes y adultos que participan en los procesos educativos.

En educación de adultos se deben tomar en consideración los aspectos propios de identidad de las personas que retoman la escolarización en programas educativos. La edad, su participación en el mundo del trabajo, sus motivaciones, sus responsabilidades civiles y sociales, diferencian a estos alumnos y alumnas de los niños y jóvenes de la escuela regular. También difieren en cuanto al tiempo de que disponen para participar en un proceso educativo, todas las modalidades de educación de adultos tienen jornadas menores y como es conocido las personas adultas disponen de poco tiempo para estudiar.

Las investigaciones sobre las características del aprendizaje en personas adultas, han mostrado la relación entre los logros que alcanzan y el grado de confianza que tienen en sí mismas. En esta relación influye poderosamente la forma en que perciben su capacidad para el aprendizaje, la que con frecuencia aparece asociada a sentimientos de desconfianza, temor y culpa, producto a veces de fracasos anteriores.

"Autoestima y autoconcepto se ligan estrechamente entre sí, influyéndose mutuamente... El autoconcepto del adulto, a diferencia de los niños, en general ya ha alcanzado una cierta estructura relativamente permanente, por lo que cada nueva experiencia educativa puede reforzarlo, fragmentarlo o dañarlo.... Ante la sensación de amenaza de sí mismo, el educando puede optar por abandonar la experiencia de aprendizaje. El adulto sólo podrá aceptar esta posibilidad de cambio y comprometerse en el proceso de aprendizaje, en la medida en que se sienta lo suficientemente seguro de que el cambio en su autoconcepto será beneficioso" (Undurraga, C. 2004).

Por ello, no basta con hacer distinciones en relación al ciclo de vida, sino también comprender las necesidades de aprendizaje y las condiciones de vida que las personas jóvenes y adultas tienen en relación a las demandas de la vida diaria.

Requiere, además, considerar que las personas "son portadoras de biografías y experiencias vitales que influyen en los estilos cognitivos; es decir, en el conjunto de características del funcionamiento mental y creativo utilizado por cada sujeto para percibir, registrar, pensar y resolver problemas" (Chilecalifica, Santiago, 2007)

Al mismo tiempo, desde el punto de vista de la evaluación, es importante recordar que los estudios que se certifican deben ser equiparables a los del sistema regular de niños y jóvenes. Esto implica que la educación de adultos también demanda exigencias de egreso, acordes con los niveles que se certifican. Para el proceso evaluativo, el reto consiste en orientar esa labor de forma tal que constituya un desafío para las personas y mida logros de aprendizaje consistentes con el nivel que se evalúa, sin que ello se convierta en un obstáculo, que desincentive la permanencia en el proceso educativo y el deseo de aprender.

2.3. Instrumentos de evaluación consistentes con las necesidades de aprendizaje de las personas adultas

Las evaluaciones que se elaboran para medir los aprendizajes de las personas adultas, deben superar una visión tradicional que la ha entendido como una educación “compensatoria” o “remedial”, respecto de lo que no pudieron realizar las personas en su niñez.

Instrumentos de evaluación que emplean textos o expresiones infantiles o que invitan a realizar tareas que “infantilizan” a las personas, distorsionan el sentido de la educación de adultos y en vez de impulsarlos a mostrar sus logros cognitivos, afectan su autoestima y limitan sus capacidades. Así, por ejemplo, cuando los adultos son invitados a realizar tareas que sólo tienen sentido en juegos de los niños o están referidas a formación de hábitos que ya adquirieron y sobre los cuales tienen juicios propios que no se recogen ni se valoran, la respuesta natural, les llevará a cohibir el desarrollo de sus pensamientos, generando una actitud de dependencia e inseguridad en el proceso educativo.

Si pensamos que la evaluación debe ser consistente con las necesidades de aprendizaje que tienen los adultos, los instrumentos que se emplean deben impulsar al despliegue de habilidades y conceptos que las personas han adquirido a través del proceso educativo y también de su experiencia.

Desde la evaluación de aprendizajes, el desafío en educación de adultos es implementar una evaluación que tenga sentido al “encajar” en la trama de vida y de intereses de las personas y sus condiciones inmediatas, pero que también despliegue expectativas y potencialidades que requieren del desarrollo de pensamiento abstracto y complejo. Ello implica que, en la evaluación se utilicen estímulos e ítems que tengan contenidos y formatos adecuados a los adultos, incorporando una perspectiva contextual, mediante el uso de temáticas generadoras que abarcan situaciones de vida de los adultos.

3. DIFICULTADES Y TENSIONES QUE ENFRENTA LA EVALUACIÓN DE APRENDIZAJE DE PERSONAS JÓVENES Y ADULTAS

Toda evaluación enfrenta desafíos técnicos y tensiones que debe resolver. Para el caso de educación de adultos hay dificultades específicas relacionadas con una escasa cultura evaluativa, con nociones y prácticas erradas en evaluación, con la heterogeneidad y complejidad de la población que participa en las ofertas educativas, con la falta de formación docente y escasa reflexión pedagógica acerca de la evaluación.

3.1. Falta de una cultura evaluativa

En educación de adultos es necesario avanzar hacia la construcción de una cultura que promueva el uso de la evaluación como punto de encuentro entre la enseñanza y el aprendizaje, que ayude – al docente y

a los alumnos (as) – a conocer los progresos y los obstáculos que tienen al momento de adquirir un nuevo conocimiento y/o desarrollar nuevas habilidades.

A esta falta de cultura evaluativa, que es transversal a todo el sistema escolar, en educación de adultos hay que agregar la persistencia de una práctica docente paternalista que tiende a compensar con buenas calificaciones a las personas por el solo hecho de permanecer en los programas. Colocar bajas exigencias de aprendizajes, tiene consecuencias, a nivel individual las personas no desarrollan todas sus capacidades y, a nivel social, se produce una deslegitimación de la certificación obtenida por alumnos y alumnas que cursan esta modalidad.

Investigaciones y observaciones directas realizadas en educación de adultos en diferentes países de América Latina, coinciden en destacar que el énfasis de la interacción está puesto en la construcción de una comunidad sobre la base del vínculo afectivo, más que de la enseñanza y logros cognitivos, se trata de “una pedagogía de la contención a expensas del conocimiento” (Brusilovsky, S. 2006)

La falta de cultura de evaluación no sólo está ausente en los profesores, sino también en los alumnos y alumnas. Como se ha señalado, las personas que participan en educación de adultos a menudo tienen una experiencia de fracaso escolar y, en el recuerdo, la evaluación es percibida como una amenaza. Estos mismos alumnos (as) que ya vivieron un fracaso anterior necesitan reafirmación constante, calificaciones insuficientes son un motivo que les puede llevar a una nueva deserción. Por su parte, es usual que un alumno(a) que no tiene tiempo ni ha logrado suficiente autonomía, le sea difícil responsabilizarse por su propio aprendizaje. Las personas buscan un reconocimiento al esfuerzo, aún cuando no perciban que el esfuerzo dedicado al aprendizaje haya sido insuficiente.

Como lo muestran investigaciones desarrolladas en el ámbito de la neurociencia “los cerebros responden mejor a ambientes complejos que a ambientes que carecen de estímulo e interés”. (OECD: 2003), de acuerdo a estas investigaciones los ambientes “con exigencias altas y pocas amenazas son aquellos que consiguen mejores resultados pedagógicos”. Así, en evaluación de aprendizajes de adultos es importante construir una cultura evaluativa que combine expectativas de logros que desafíen a los alumnos y que, al mismo tiempo, no resulten amenazantes

3.2. Diversidad de la población: motivación y expectativas

Como ha sido señalado un rasgo que caracteriza a la educación de jóvenes y adultos es la diversidad de programas y la heterogeneidad de la población que acude a los mismos. La diversidad se puede enfocar desde distintos aspectos: la escolaridad, la edad, la identidad sociocultural. Todos estos aspectos inciden también en otro: la motivación que los impulsa para volver a estudiar.

En relación con la edad, es oportuno señalar que todas las estadísticas y estudios muestran que la población que acude a la modalidad escolar de adultos está conformada fundamentalmente por jóvenes. En el informe regional preparado para CONFITEA VI, su autora describe “hubo un tiempo en que los jóvenes tenían que adaptarse a los programas y clases de educación de adultos. Ahora que los jóvenes han sido ubicados en el centro de la educación de personas jóvenes y adultas (EPJA), esto está llevando a una discriminación de los adultos dentro de la EPJA” (Torres R.M. 2009).

A modo de ejemplo, en Chile, la educación tradicional de adultos que se realiza en Terceras Jornadas y Centros Integrales de Educación de Adultos (CEIA's) el promedio de edad es de 28 años para la Educación Básica y 25 años para la Educación Media. En cambio, en los programas flexibles de nivelación de estudios estos promedios se elevan en 38 y 33 años respectivamente.

En educación de adultos también se debe considerar la inclusión de poblaciones que tienen rasgos específicos ya sea por identidad cultural, como es el caso de las etnias; ya sea por distinción entre géneros; o bien por ubicación temporal, como es el caso de los internos en recintos penales o bien de poblaciones rurales apartadas.

Ocurre así, que un mismo currículum puede ser aplicado y conducir a una misma certificación de estudio, en poblaciones que están en diferente ciclo de vida y tienen identidades sociales y culturales específicas, para los cuales el proceso educativo tiene un significado distinto y que difieren también en cuanto a expectativas y motivaciones en relación con los estudios.

Si consideramos esta pluralidad de realidades y diversidad de personas que se incorporan a la educación de adultos, desde la evaluación se plantean importantes desafíos: ¿es posible establecer los mismos parámetros de evaluación para toda la población?, ¿en sistemas nacionales de evaluación que conducen a certificación de estudios, cómo hacer para construir instrumentos válidos, que considere esta pluralidad y controle posibles sesgos?, ¿el docente en el aula puede y/o debe colocar las mismas exigencias de logros a personas que están en etapas diferentes del ciclo de vida?.

Si para todo proceso de evaluación que conduce a la certificación, uno de los desafíos más importantes es construir instrumentos válidos - que den cuenta efectiva de los aprendizajes alcanzados por los alumnos y alumnas, de acuerdo al currículum vigente y según el nivel escolar que se examine - para el caso de la educación de personas jóvenes y adultos, este desafío se torna aún más complejo.

3.3. La construcción de conocimientos a partir de la experiencia y obstáculos cognitivos

Los adultos y los jóvenes que buscan una segunda oportunidad educativa construyen, mucho más que los niños, el conocimiento a partir de su experiencia. En el proceso de aprendizaje las personas adultas actúan desde su contexto, desde sus experiencias, usando las estrategias que ya han desarrollado en otras ocasiones, las que asimilan, por alguna razón de semejanza, al proceso que se ven envueltas o al problema que intentan resolver o entender.

“Cuando se enfrenta un nuevo contenido a aprender, lo hace siempre armado de un bagaje de conceptos y representaciones adquiridos en el transcurso de sus experiencias previas, las que utiliza como instrumento de lectura o interpretación y que determinarán en buena medida qué informaciones seleccionará, cómo las organizará y qué tipo de relaciones establecerá entre ellas” (Coll, 1990)

Desde este enfoque, el aprendizaje se torna significativo cuando es posible relacionar lo nuevo con lo que se conoce y se sabe; lo que va a aprender con lo ya aprendido. Para que el nuevo conocimiento posea consistencia, su construcción necesita cimentarse sobre la base de conocimientos ya existentes.

Sin embargo, el énfasis en los aprendizajes y experiencias previas como vínculo con los nuevos conocimientos requiere ser relativizado, “la relación de los nuevos contenidos con los esquemas mentales previos de los educandos no es tan directa y armónica: el aprendizaje necesita pasar a veces por un preliminar “*desaprendizaje*” de nociones erróneas”. (Kaplún, M. 1995).

De esta manera, es necesario considerar que en las experiencias de vida existen obstáculos cognitivos - convicciones erróneas- que tienen un estatus de verdad en las personas y que bloquean la adquisición de nuevos conocimientos (Cfr. Astolfi, 2003). Por ello, las personas jóvenes y adultas requieren de un proceso educativo que sea capaz de ayudarlos a superar esos obstáculos.

El resultado de evaluaciones de aprendizaje muestra que la experiencia de las personas puede constituir un poderoso punto para el aprendizaje; sin embargo en muchas ocasiones el pensamiento común también puede constituir un obstáculo. A lo largo de su vida las personas han construido explicaciones que poseen una coherencia indiscutible, aunque sean “científicamente falsas”. Así, muchas veces el docente en el aula no percibe con claridad la construcción de razonamientos paralelos y sólo al momento de la evaluación, estos razonamientos se hacen evidentes.

Cuando a partir de la experiencia directa se construyen y/o validan los conocimientos, no hay rupturas cognitivas, ya que es fácil aceptar lo que se observa, palpa o siente. No sucede lo mismo cuando se quiere incorporar conceptos que van en contra del “sentido común”. Así, por ejemplo³, cuando se solicita mencionar los estados del agua (sólido, líquido y gaseoso) es habitual que los adultos reconozcan como agua sólo la que está en estado líquido. Aunque parezca elemental, las personas, aún en niveles escolares avanzados, espontáneamente, rechazan la idea de pensar que el agua puede estar en estado sólido. Cuando se busca apoyar el razonamiento a partir de lo que se puede observar, las personas suelen expresar: *“ah, pero es que eso es hielo, no es agua”*. La experiencia propia y los conceptos apropiados desde la vida cotidiana son más fuertes que afirmaciones realizadas en un contexto escolar.

Otras situaciones muestran como, a veces, el “sentido común” y la experiencia inmediata genera asociaciones erradas que impiden contestar de acuerdo a conceptos disciplinares que son requeridos en el currículo. Así, por ejemplo, luego de la lectura de un texto acerca de las relaciones sociales y económicas correspondientes al período colonial, a los alumnos se les solicitó que señalaran alguna característica del sistema de encomienda; muchos adultos se refirieron a este concepto relacionándolo con el envío de un paquete a través del correo.

Otro ejemplo que permite observar cómo se impone la experiencia directa de las personas se puede verificar con relación a una tarea que mide la identificación de información explícita en un texto breve. Como estímulo se utilizó una receta para cocinar “charquicán” se solicitaba a las personas contestar de acuerdo a la información y, sin embargo, un número apreciable respondió a partir de la manera habitual en que cocina la receta obviando la información del texto.

Estos ejemplos muestran como la información se lee desde lo que se conoce y, muchas veces, especialmente en los lectores incipientes, no hay plasticidad suficiente para adecuarla a una nueva situación. La evaluación puede ser una poderosa herramienta para detectar convicciones erróneas, en este sentido, el educador puede cumplir un papel como investigador permanente de las estrategias que desarrolla su grupo en la resolución de diferentes tareas, descubriendo las asociaciones que realizan y la forma concreta en que logran los resultados, sean estos los considerados correctos o no.

3.4. Falta de formación docente y de condiciones para planificar la evaluación

En América Latina es conocida la falta de formación inicial de educadores de adultos. Existen pocos centros de formación y sólo unas pocas facultades de educación la incorporan como especialización profesional. Con mayor frecuencia se encuentran ofertas de formación dirigida a los docentes en servicio. Estos cursos suelen ser de corta duración y, en estos, la evaluación de aprendizaje no constituye una temática prioritaria.

³ Ejemplos escogidos de experiencias de validación de ítems de evaluación. Línea de Evaluación de Aprendizaje y Certificación de Estudios. Sistema de Formación Permanente Chilecalifica.

Junto a la falta de formación específica, es necesario considerar los contextos laborales de los docentes: inestabilidad laboral, horarios parciales, trabajo simultáneo en varios establecimientos educativos y/o en programas flexibles de nivelación de estudios, combinación de la labor docente con otros trabajos para mejorar los ingresos.

Por otro lado, como ha sido observado, “es habitual verificar una situación de abandono que experimenta la enseñanza para adultos en los establecimientos. El sistema es débil, se trata de contextos en los cuales los profesores y profesoras tienen total libertad de acción, pero a la vez no tienen el apoyo de un proyecto institucional con una organización clara”⁴.

Otro factor que los docentes señalan como característico en sus contextos, es la irregularidad en la asistencia de los estudiantes y la diversidad de niveles iniciales que tienen los alumnos(as), muchos de los cuales no han alcanzado las competencias que acreditan haber obtenido. Estas situaciones contextuales, verificada en la educación escolar de adultos, acaba por incidir en la posibilidad efectiva de planificación de la enseñanza y evaluación de los aprendizajes en el aula.

Así, a la dificultad habitual de los docentes de contar con tiempo y formación suficiente para su propia formación; en el caso de educación de adultos, se suma la escasa oferta de profesionalización docente, la frecuente rotación de alumnos e irregularidad en la asistencia a clases y una débil institucionalidad que obstaculiza la planificación del trabajo en el aula.

En el ámbito internacional, el reconocimiento de esta realidad se expresa en las recomendaciones establecidas en la reunión regional de CONFITEA, al señalar la necesidad de “elaborar políticas de formación inicial y permanente de educadores de personas jóvenes y adultas con la participación de las universidades, los sistemas de enseñanza y los movimientos sociales para elevar la calidad de los procesos educativos y asegurar el mejoramiento de las condiciones laborales y profesionales de los educadores y funcionarios” (CONFITEA VI, México:2008)

4. ¿CÓMO AVANZAR EN LA EVALUACIÓN DE APRENDIZAJES DE PERSONAS JÓVENES Y ADULTAS?

Si bien no se pueden proponer modelos únicos para acciones educativas tan diversas y complejas, sí se pueden identificar algunas claves para avanzar en la evaluación de aprendizajes de personas jóvenes y adultas, entre ellas: el diseño de evaluaciones que amplíen el enfoque tradicional e incorporen la perspectiva de medición de competencias; el diseño y elaboración de instrumentos basados en un enfoque criterial o edumétrico, y el diseño y construcción de sistemas de evaluación para la acreditación, en el contexto de las nuevas políticas de educación de adultos.

4.1. Evaluación en la perspectiva de las competencias

En educación de personas jóvenes y adultas existe la necesidad de ampliar el enfoque de evaluación, evitando una concepción basada en la reproducción de conocimientos, para incorporar una perspectiva de evaluación por competencias.

⁴ Extraído de informe elaborado por equipo de perfeccionamiento docente de la Universidad de Chile: 2007

La amplia bibliografía sobre el enfoque por competencias en educación, coincide en señalarlo como un concepto integrador en la medida que considera tanto los contenidos, como las habilidades puestas en práctica en diferentes situaciones de vida. De acuerdo con este enfoque, la competencia se entiende como “la capacidad de actuar de manera eficaz en un tipo definido de situación, capacidad que se apoya en conocimientos pero no se reduce a ellos” (Perrenoud, P. 1999). Con esta misma orientación se definen las competencias “como un conjunto ordenado de capacidades que se ejercen sobre contenidos en una categoría de situaciones para resolver problemas que ellas generan” (De Ketele, 1996).

Al incorporar este enfoque en las evaluaciones se buscará medir conocimientos – entendidos como los procesos de comprensión, apropiación y aplicación de saberes conceptuales – y habilidades – entendidas como procedimientos mentales que se aplican en la práctica y ayudan a movilizar nuevas adquisiciones cognitivas – puestas en situaciones en las cuales las personas deben movilizar estos recursos.

Incorporar el enfoque por competencias implica reconocer que el aprendizaje se logra en diferentes ámbitos de la vida, no sólo en el espacio formal curricular, sino también a través de la experiencia y de la interacción en espacios no formales e incluso informales; implica, además, reconocer que en la vida diaria las personas al enfrentar una situación desafiante deben utilizar y asociar variados recursos cognitivos complementarios, entre los cuales se encuentran los conocimientos y habilidades.

Así como en la vida cotidiana, y ante situaciones desafiante, las personas ponen en actividad sus conocimientos y habilidades, en la evaluación basada en competencias se trata de que las personas “movilicen” sus capacidades cognitivas, de modo que les permitan evocar, organizar, seleccionar y jerarquizar en función de una tarea específica solicitada en la evaluación.

El enfoque por competencias ubica los contextos como el escenario donde los conocimientos se aplican, verifican y cobran sentido. El contexto se refiere a un entorno específico, a un “entramado” que da sentido y significado al pensamiento, a las actitudes, a los valores. Por tanto, el contexto no remite solo a un espacio geográfico o ambiental, sino a una trama compleja que incluye el espacio y la situación que allí ocurre. La familia, el barrio o el trabajo pueden constituir un contexto, así como el país o el ámbito internacional pueden formar parte también de la experiencia vital de las personas.

Así, en el diseño de instrumentos que asumen esta perspectiva de evaluación, se debe buscar que las tareas solicitadas se presenten en situaciones que varíen en complejidad e intención, según el propósito del programa y, para el caso de las pruebas de certificación, según el nivel escolar que se evalúa.

Coherente con los planteamientos de este enfoque, se debe intencionar que en los instrumentos de evaluación se presenten situaciones que trasciendan al ámbito escolar. Por tal motivo, en la elaboración de los mismos, será necesario que los ítems estén relacionados con estímulos de base presentados a través de textos de diferente tipo y presentado en diferentes formatos, ilustraciones, noticias, gráficos, tablas, etc. que presenten situaciones referidas a un contexto.

4.2. Diseño y elaboración de instrumentos basados en un enfoque criterial o edumétrico

La evaluación, ya sea la que desarrolla el docente en el aula o la que se implementa a través de sistemas de acreditación, debe ser un medio para mejorar el aprendizaje, de retroalimentación a los alumnos y a los propios docentes. Ello requiere establecer un marco y criterios de logros transparentes y claros para alumnos(as) y docentes, de manera que las personas sepan qué se espera de ellas y qué tienen que hacer. En este ámbito, en la educación escolar de adultos, se plantea el desafío de construir marcos que

conformen una selección del curriculum con fines evaluativos, a partir de los cuales se avance hacia la definición de estándares.

En el enfoque criterial o edumétrico se trata de avanzar hacia una evaluación que supere el enfoque tradicional basado en normas – que tienen como propósito principal ordenar comparativamente a los estudiantes – a pruebas que se enfocan en la definición de qué es lo que las personas saben y son capaces de hacer. Es decir, se trata de definir cuál es el resultado que los alumnos y alumnas deberían alcanzar para que su desempeño pueda ser considerado como satisfactorio.

En educación de personas jóvenes y adultas, existen mejores condiciones para consensuar y reflexionar acerca de lo que se desea alcanzar en el proceso educativo, si las personas son informadas se sienten participantes de su proceso, se comprometen más.

Por otro lado, los criterios e indicadores de evaluación deben comunicarse de manera adecuada para sean comprendidos por docentes y alumnos, esta es la base de una evaluación entendida como retroalimentación para que las personas sepan en qué han fallado y qué les falta por saber. Desde esta perspectiva se podrá superar la idea de la evaluación como la obtención de una calificación, que sólo comunica una nota, pero no permite transparentar el proceso de enseñanza y aprendizaje.

En educación de adultos se requiere avanzar hacia la definición de niveles de logros que señalen con claridad no solo los contenidos y objetivos que son evaluados, sino cuál es el nivel de profundidad con que ellos debieran ser alcanzados. Este es un desafío de futuro, no es fácil llegar a establecer cuáles son las competencias que debieran lograr las personas y de acuerdo a qué estándares.

Avanzar hacia la definición de criterios y de estándares de evaluación, exige simultáneamente generar oportunidades de aprendizaje. En educación de adultos existe urgencia de contar con los medios que permitan a las personas y las instituciones asegurar que todos ellos tengan acceso a condiciones y oportunidades educativas, para alcanzar logros de aprendizaje de acuerdo a estándares de calidad.

4.3. La construcción de políticas para el desarrollo de sistemas de evaluación y acreditación

La convicción de que las competencias se aprenden y desarrollan en los diferentes ámbitos educativos: formal, no formal o informal, plantea el desafío de construir sistemas de evaluación que acrediten tanto los aprendizajes desarrollados por las personas en el sistema escolar como los adquiridos a través de la experiencia. Contar con sistemas de evaluación permite la construcción de ofertas educativas flexibles y facilita que las personas puedan generar sus propias trayectorias educativas, teniendo como enfoque el aprendizaje a lo largo de la vida y su articulación con sistemas de formación permanente.

“La educación permanente genera a las naciones el desafío de encontrar nuevas formas de comprobación y certificación de los aprendizajes que la competitividad social exige. En las instancias de aprendizajes formales de ello se ocupa el sistema educativo. Pero la educación a lo largo de toda la vida presupone el desarrollo de nuevas formas de certificación en las que se consideren todas las competencias logradas, independientemente de los procesos de formación seguidos” (Picaroni, B. 2008).

Construir políticas de acreditación de los aprendizajes logrados requiere, simultáneamente, que los Estados se hagan cargo de impulsar ofertas educativas que permitan el logro de estos aprendizajes. El discurso del aprendizaje permanente no puede ser entendido como responsabilidad del individuo de su propia trayectoria o proceso, como tampoco las oportunidades que sistemas de evaluación otorgan para acreditar estudios, puede reemplazar el deber del Estado y el derecho ciudadano de brindar y recibir educación. De esta manera, la construcción de sistemas de evaluación y acreditación deben enmarcarse

en políticas más amplias de educación de personas jóvenes y adultos y, su impulso, se debe relacionar con la necesidad de ampliar y cautelar la calidad de la oferta educativa.

En el ámbito de evaluaciones a gran escala y externas al proceso educativo, la experiencia acumulada está fundamentalmente referida a las evaluaciones cuyo propósito es medir la calidad de la educación de los niños y jóvenes del sistema regular, sin consecuencias individuales para la aprobación de cursos o niveles educativos.

Desde educación de adultos, se están desarrollando experiencias pioneras de construcción de sistemas cuyo propósito es la acreditación. Estos sistemas tienen impacto en dos dimensiones: a nivel político respecto de la transparencia, validez y confiabilidad de sus resultados y, a nivel individual, con impacto directo en la certificación de los estudios para las personas.

"Un sistema de evaluación para la certificación de los aprendizajes de los estudiantes a través de exámenes nacionales de alta calidad técnica tiene las siguientes ventajas. Hace transparente el valor del certificado; en el presente es común que dos estudiantes que han logrado aprobar un mismo nivel educativo en dos regiones distintas de un mismo país, en realidad posean niveles de conocimiento muy diferentes. Hace responsables tanto a los docentes como a los propios estudiantes por alcanzar los conocimientos y capacidades que son objeto de evaluación en el examen, lo cual tiene impactos positivos en el aprendizaje." (Ravela, P. 2007)

Un sistema de evaluación de alta calidad técnica requiere de un conjunto de definiciones teóricas y procedimentales relacionadas con:

- el propósito, alcance y enfoque de la evaluación;
- el levantamiento de marcos de evaluación que orienten la selección de los contenidos y objetivos del curriculum que serán medidos;
- el diseño de instrumentos;
- el formato y tipo de ítems que se aplicarán;
- la construcción de criterios, indicadores y pautas de corrección para cada ítems;
- el diseño de metodologías de validación y análisis de los instrumentos;
- la definición de puntos de corte;
- la determinación de procedimientos de aplicación y corrección;
- el diseño de estrategias de información de los resultados.

Para el caso de las pruebas de certificación es especialmente relevante la descripción de niveles de logro que señalen con claridad no sólo los contenidos y objetivos que son evaluados, sino cuál es el nivel de profundidad con que ellos debieran ser alcanzados en determinado sector y nivel escolar. A nivel técnico, ello implica contar con procedimientos de análisis estadísticos y cualitativos de ítems e incorporar metodologías para definir puntos de corte.

Un sistema de evaluación requiere de capacidades técnicas y equipos profesionales que puedan desarrollar los procedimientos señalados y cuenten con capacidad para gestionar las diferentes etapas del proceso. Estos sistemas son complejos, son procesos factibles de construir de manera gradual y para su maduración se necesita de un tiempo prolongado y de recursos que permitan su desarrollo.

También es necesario anticiparse a las tensiones que un sistema de acreditación dirigido a las personas jóvenes y adultas debe enfrentar⁵, la más importante de todas consiste en calibrar adecuadamente el nivel de exigencia: si las pruebas son muy exigentes, puede llevar a un alto nivel de reprobación, por el contrario, si las pruebas se elaboran pensando en la aprobación, se envían señales equivocadas acerca de las expectativas de aprendizaje y se invalida el propio proceso de certificación⁶.

Los sistemas de certificación requieren otorgar varias oportunidades de examinación y brindar a las personas la posibilidad de acreditar parcialmente sectores de aprendizaje, así puedan avanzar en forma paulatina hasta obtener la certificación de estudios del nivel escolar correspondiente. Ello plantea tres grandes tensiones: la demanda permanente de construcción y validación de ítems que permita ensamblar pruebas equiparables entre sí; la necesidad de implementar varios operativos de aplicación en el año controlando la transparencia del proceso y confidencialidad de los instrumentos; la exigencia de informar fluida y claramente de los resultados después de cada aplicación, buscando comunicar a los docentes y alumnos (as) las tareas que fueron evaluadas y los logros obtenidos por las personas examinadas.

Derivado de lo anterior, otra tensión y desafío constante consiste en controlar la confiabilidad y confidencialidad en la aplicación y corrección de los instrumentos, buscando garantizar condiciones homogéneas a todas las personas que se presentan a examen.

Por último, es importante establecer los límites de un sistema estandarizado que tiene como propósito la acreditación. Por sus características, generalmente, sólo pueden aplicar pruebas de "lápiz y papel" y en este tipo de pruebas es difícil incorporar ítems de desarrollo que permitirían de mejor manera verificar el proceso que ha seguido la persona para resolver una situación, o bien medir la capacidad de formular un argumento, justificar sus decisiones, sus puntos de vistas o producir textos escritos. De igual modo, es muy difícil, en este tipo de pruebas, medir objetivos valiosos de la educación como la formación de actitudes y comportamientos.

REFERENCIAS BIBLIOGRÁFICAS

- Astolfi, J.P. (2003). *Aprender en la escuela*. Santiago de Chile: LOM.
- Coll, C. (1990) Un marco de referencia psicológico para la educación escolar: la concepción constructivista del aprendizaje y la enseñanza. En Coll, C. Palacios, J. y Marchesi, A. (comp.) *Desarrollo psicológico y educación*. Tomo II. Alianza: Madrid.
- CHILECALIFICA. Línea de Evaluación y certificación de estudios (2007). *Evaluación de aprendizajes desde el enfoque de competencias*. Santiago de Chile: CHILECALIFICA.
- CHILECALIFICA. Línea de Evaluación y certificación de estudios (2008). *Sistema Nacional de Evaluación y Certificación de Estudios para personas jóvenes y adultas*. Documento elaborado para el Seminario

⁵ Reflexiones basadas en la experiencia de seis años de construcción del Sistema Nacional de Evaluación de Aprendizajes desarrollado en Chile, por el Ministerio de Educación y el Programa Chilecalifica.

⁶ En sistemas consolidados como es el caso del General Educational Development Testing Program (GED) de Estados Unidos, la definición de los estándares requeridos para la aprobación de las cinco áreas de contenido correspondiente a la educación secundaria, lo realiza un comité de especificaciones de la prueba conformado por profesores y expertos de educación secundaria, que se reúnen cuatro o cinco años antes de empezar a armar la prueba.

- Internacional de Evaluación de Aprendizajes, Santiago de Chile: CHILECALIFICA.
- De Ketele, J.M. (1996). L'évaluation des caquis scolaires: quais? Por quoi?. *Revue des Sciences de l'Education*.
- Fundación Paulo Montenegro, Acao Educativa.(2007). *Indicador de Alfabetismo Funcional*. INAF http://www.ipm.org.br/ipmb_pagina
- Infante, M.I. (2000). *Alfabetismo funcional en siete países de América Latina*. Santiago: UNESCO-OREALC.
- Kaplún, M. (1995). *Los materiales de autoaprendizaje: marco para su elaboración*. Santiago de Chile: UNESCO.
- Kirsch I. (2001). *The international Adult Literacy Survey (IALS): understanding what was measured*. Princeton: Education Testing Service.
- OCDE. (2003). *La comprensión del cerebro. Hacia una nueva ciencia del aprendizaje*. Madrid: Santillana
- Perrenoud, P. (1999). *Construir competencias desde la escuela*. Santiago de Chile: Dolmen Pedagogía.
- Picaroni, Beatriz (2008). Desafíos para la construcción de un sistema destinado a la evaluación de aprendizajes y certificación de estudios. Ponencia presentada en el *Seminario Internacional de Evaluación de Aprendizaje de Personas Jóvenes y Adultas*.
- Ravela,P. et al. (2007). *Las evaluaciones educativas que América Latina necesita*. Santiago de Chile: PREAL.
- Secretaria de Educação Continuada, Alfabetização e Diversidade (SECAD) /UNESCO/IPEA(2006). *Brasil Alfabetizado: Caminhos de avaliação. Coleção Educação para todos*. Brasília:SECAD. <http://unesdoc.unesco.org/images/0015/001545/154573por.pdf>
- Sistema Automatizado de Seguimiento y Acreditación, México (2007) http://www.conevyt.org.mx/cursos/cursos/asesor/mainframe/folletos/folleto1/tema_01.htm
- Torres, R.M. (2009). De la alfabetización al aprendizaje a lo largo de toda la vida. Informe Regional, Preparado para la *VI Conferencia Regional sobre Educación de Adultos*.
- UNESCO/UII (2008) *Compromiso renovado para el aprendizaje a lo largo de toda la vida. Propuesta de la región de América Latina y El Caribe*. Santiago de Chile: UNESCO.
- Undurraga C. (2004). *¿Cómo Aprenden los Adultos?*. Santiago de Chile: Ediciones Universidad Católica de Chile.

EL ANÁLISIS COGNITIVO DE TAREAS, BASE PARA EL DISEÑO DE INSTRUMENTOS DE EVALUACIÓN EN EL APRENDIZAJE EN LÍNEA

Eduardo Peñalosa y Sandra Castañeda

Revista Iberoamericana de Evaluación Educativa 2009 - Volumen 2, Número 1

<http://www.rinace.net/riee/numeros/vol2-num1/art9.pdf>

Fecha de recepción: 30 de marzo de 2009

Fecha de comunicación de dictamen: 29 de abril de 2009

Fecha de aceptación: 29 de abril de 2009

Se estima que el aprendizaje en línea es un proceso activo y complejo, en el que el estudiante construye sus conocimientos con base en una estructura de significados previos, mediante la interacción con agentes humanos o materiales en entornos virtuales. Para el desempeño de estos procesos se ha sugerido que los ambientes de aprendizaje en línea estén centrados en: a) el estudiante como protagonista y agente; b) el conocimiento y la competencia; c) la evaluación, para ayudar a los estudiantes a tener una valoración de su propia comprensión, y d) la comunidad, que propician la colaboración con otros estudiantes y miembros de la misma (Bransford, Brown y Cockings, 2004; Anderson, 2004).

De entre las funciones enunciadas antes, en este trabajo interesa en especial la de evaluación del desempeño en contextos de aprendizaje en línea. Se describe la construcción de evaluaciones sensibles del aprendizaje, que permiten identificar la influencia de elementos del ambiente y del diseño instruccional. Para identificar el progreso de los procesos cognitivos e interactivos en entornos virtuales de aprendizaje es necesario contar con una estrategia sensible y válida de evaluación del desempeño, así como una serie de herramientas que permitan detectar cambios en la complejidad de las construcciones de conocimientos por parte de los estudiantes.

Existen tres etapas en la evaluación del aprendizaje presencial o en línea: 1) la que se realiza inicialmente de manera que tengamos una idea fiable del nivel de conocimientos y habilidades de los estudiantes (evaluación inicial); 2) la que se realiza durante la acción docente, a intervalos establecidos ya en la fase de diseño y planificación (evaluación formativa), y 3) la que se realiza al finalizar la acción docente (evaluación sumativa).

Dadas las características de la educación en línea, donde existen limitaciones en la comunicación impuestas por la naturaleza del medio, la evaluación continua del aprendizaje es esencial, por su función de retroalimentación al estudiante acerca de su desempeño. Quesada (2006) plantea que los atributos de una evaluación en línea deben ser: confiabilidad, validez, objetividad y autenticidad. La primera se refiere a la confianza que genera un instrumento para reflejar el nivel de logro del estudiante; la segunda se refiere a que el instrumento mida lo que realmente se pretende, y no otra cosa; la tercera se refiere a la neutralidad con la que se califica a los estudiantes, y la última se relaciona con los contenidos de la evaluación con la realidad práctica de los estudiantes. Quesada indica que si bien se ha escrito mucho sobre educación a distancia "en línea", pocas referencias se dedican al análisis de la evaluación en estos contextos.

En la investigación acerca del aprendizaje en línea, Peñalosa y Castañeda (2008a) citan alrededor de una decena de estudios en los que se analiza la eficacia de cursos en línea de Psicología, y las estrategias de evaluación son, fundamentalmente: 1) la realización de cuestionarios de conocimientos que se presentaban al finalizar los cursos, con un dudoso cumplimiento de criterios de confiabilidad, validez o autenticidad, aunque con cierta garantía de objetividad, inherente al medio en línea, y/o 2) la aplicación de entrevistas o inventarios que sondean la satisfacción del estudiante en el curso, lo cual implica considerar como iguales satisfacción y aprendizaje, dos aspectos no necesariamente equiparables, ya que la variable satisfacción hace mención más a efectos motivacionales que instruccionales.

La variedad de recursos de evaluación del aprendizaje en línea incluye pruebas objetivas, ensayos, proyectos, rúbricas, creación de mapas conceptuales o evaluación basada en portafolios de evidencias, por citar las más utilizadas. Una clasificación posible de los tipos de evaluación en la educación en línea incluye evaluaciones: a) automáticas, programadas para ser resueltas y calificadas de inmediato; b) elaborativas, en ellas el estudiante construye un producto como un ensayo, un mapa conceptual, un

proyecto, etc.; y c) colaborativas, en ellas los estudiantes son evaluados en situaciones de trabajo grupal (clasificación adaptada de Barberá, 2006).

Un paso crítico en el desarrollo de sistemas centrados en el estudiante consiste en tener información de calidad acerca de las competencias del estudiante en el dominio educativo. La tecnología incluye un conjunto de herramientas con potencialidad para optimizar la evaluación de los aprendizajes en entornos en línea. De hecho, en la modalidad educativa en línea, la evaluación debería ser una actividad continua, el eje vertebrador del aprendizaje (Barberá, 2006).

Sin embargo, existe coincidencia entre algunos autores acerca de que si bien en la puesta en marcha de proyectos de educación virtual se ha prestado atención a la pertinencia de los contenidos, al diseño instruccional, al uso óptimo de los recursos tecnológicos, por otra parte se ha dejado de lado la consideración minuciosa de la evaluación del aprendizaje que con ellos se logra (Quesada, 2006), y de hecho la evaluación se ha considerado un “punto débil” de la educación virtual (Barberá, 2006).

La infraestructura tecnológica podría permitir el registro flexible y exhaustivo de los datos de desempeño y el seguimiento de los progresos de los alumnos en su proceso de desarrollo cognitivo, de construcción de conocimientos y adquisición de habilidades. En estos contextos el sistema sigue la huella de lo que el estudiante hace, no omite partes de su desempeño y tampoco emite juicios parciales: tiene la programación de los componentes que se deben cumplir y registra los datos en consecuencia. Una evaluación tan minuciosa sería difícil de llevar a la práctica por parte de un profesor, y por su minuciosidad se podría prestar a omisiones. Para un curso determinado, se programa una vez y se utiliza para cientos de alumnos. Adicionalmente, se convierte en un laboratorio para identificar las mejores prácticas instruccionales y los efectos de los distintos tipos de esquemas interactivos.

A pesar de tal potencialidad, la evaluación del aprendizaje en línea es “algo pendiente de mejorar, algo sobre lo que hay que reflexionar con mayor profundidad, que hay que desarrollar de un modo más creativo y volver sobre ello invirtiendo más recursos pedagógicos y tecnológicos” (Barberá, 2006: 4).

De esta manera, tenemos que si bien la evaluación se considera parte fundamental del proceso de aprendizaje en línea, y aun cuando se reconoce un gran potencial de la tecnología para crear sistemas efectivos de evaluación, no existe un desarrollo relevante de aplicaciones tecnológicas para realizarla.

La propuesta de este trabajo se desarrolla en el dominio de un curso de psicología clínica teórica. Dada la necesidad de contar con este tipo de evaluaciones que brinden información acerca de los niveles de desempeño en los que impacta la intervención, en el presente estudio se reporta la metodología para construir y validar instrumentos de evaluación para la enseñanza en línea de contenidos de psicología clínica teórica, así como los resultados de dicha validación.

Dada la identificación de métodos instruccionales tradicionales, que conducen a un aprendizaje superficial o inerte, que permite a los estudiantes recordar o repetir lo aprendido, pero no utilizar el conocimiento nuevo como herramienta; esto es, no demostrar la capacidad de aplicación ni transferencia del mismo (Renkl, Mandl y Gruber, 1996; Schwartz, Brophy, Lin y Bransford, 1999), se hace necesario contar con herramientas diagnósticas que permitan considerar los diversos niveles de complejidad de los procesos subyacentes al desempeño, así como evaluar integralmente la progresión del estudiante en las diversas etapas y capas del aprendizaje. Esto es posible con ayuda de un análisis diferencial de operaciones cognitivas de complejidad creciente; esto es, con un *Análisis Cognitivo de Tareas (ACT)* (Castañeda, 1998a; 1998b; 2002; 2004a; 2004b; 2006; Chipman, Schragen y Shalin, 2000).

1. ANÁLISIS COGNITIVO DE TAREAS Y EVALUACIÓN DEL DESEMPEÑO

Para realizar un ACT, en primera instancia es necesario identificar las competencias que componen al dominio educativo. Podemos plantear que esta identificación de las competencias principales de un dominio constituye la macroestructura de lo que se desea evaluar (Castañeda, 2006). Cuando contamos con ella, se procede a realizar un ACT para identificar los microcomponentes que constituyen a cada competencia, e identificar gradientes de complejidad basados en las demandas de procesamiento generadas en la situación de evaluación. El procedimiento analiza la tarea mediante pasos en los que se identifican, en una secuencia progresiva de mayor detalle y precisión, los conocimientos, habilidades y las disposiciones asociados a cada paso. En este proceso es importante identificar la complejidad cognitiva (p. ej.: comprensión, aplicación, solución de problemas) y el tipo de modelo mental (p. ej.; conceptual, estructural, causal).

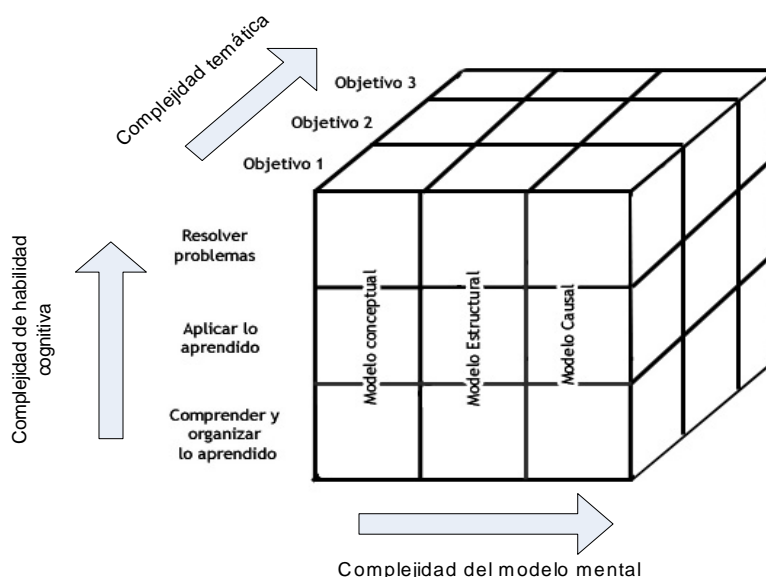
Con base en ambos análisis, se construye una estructura que identifica componentes de grano fino sin perder de vista las ligas de las competencias de las que forman parte. De esta forma se organiza un universo de medida tal que ha permitido diseñar situaciones de evaluación representativas de competencias críticas y sus componentes, en un esquema articulado y anidado de evaluación.

Como se expone arriba, en el proceso de análisis de componentes de las tareas es preciso clasificar los pasos del desarrollo de las competencias, y de esta forma se obtienen gradientes de complejidad de las competencias. Las dimensiones para realizar estas clasificaciones son:

1. Un continuo de complejidad en cuanto a las habilidades cognitivas requeridas en el dominio, que incluye tres categorías:
 - a) comprensión de los temas, que incluye el reconocimiento de información, la clasificación, el ordenamiento, el resumen;
 - b) aplicación del conocimiento y la habilidad, que incluyen operaciones como la traducción, el análisis, la extrapolación, la inferencia, la comparación o la aplicación de procedimientos, y,
 - c) solución de problemas, que incluye operaciones como corrección de errores, planificación de acciones, evaluación y toma de decisiones.
2. La dimensión de complejidad en cuanto a los modelos mentales construidos por los estudiantes. Las habilidades cognitivas representan categorías de operaciones que el estudiante puede aplicar en el dominio; los modelos mentales reflejan la integración de conocimiento que le permiten explicar la realidad de los fenómenos. La inclusión de los modelos mentales como un eje de diagnóstico responde a evidencias de que los estudiantes, en función de su nivel de pericia, podrían describir sólo conceptualmente el dominio, o podrían tener un conocimiento altamente estructurado acerca del mismo, por lo cual podría explicarlo. Los modelos mentales que se incluyen en el presente modelo son tres:
 - a) conceptuales, responden a la pregunta: "¿qué es esto?", describen el significado del fenómeno o tema, y la interrelación de los elementos que lo componen;
 - b) estructurales, que responden a la pregunta: "¿cómo está estructurado esto?", y describen cómo está organizado el campo conceptual en cuestión; y,

- c) causales, que responden a la pregunta: "¿cómo funciona esto?", y que describen cómo los principios se afectan entre sí y ayudan a interpretar procesos, dar explicaciones de eventos y realizar predicciones.
3. La dimensión de la complejidad de los temas de aprendizaje, determinada por las unidades de contenido como universo a evaluar aplicando diferentes modelos mentales con diferentes niveles de complejidad cognitiva. Esta dimensión se ordena de lo simple a lo complejo, en función de los objetivos de las unidades del programa de aprendizaje elegido.

FIGURA 1. MODELO DEL ANÁLISIS COGNITIVO DE TAREAS DE UN DOMINIO



La figura 1 muestra la estructura que resultaría de integrar las tres dimensiones relevantes del análisis. En ella cada espacio del cubo incluye una categoría de contenido que juntas conforman el universo del dominio educativo analizado. Esto significa que de cada casilla pueden generarse reactivos de evaluación, recursos y actividades de aprendizaje.

Por ejemplo, tomemos el dominio educativo de la evaluación en psicología clínica. La casilla inferior izquierda tendría contenidos del primer objetivo (digamos antecedentes teóricos de la evaluación psicológica), del nivel cognitivo de comprensión (implica identificar definiciones de evaluación, de evaluación como signo, como muestra), y del modelo mental conceptual (el estudiante debería encontrar instancias de tareas de evaluación psicológica entre varias disponibles).

En el mismo dominio, una casilla que implica mayor complejidad: imaginemos la superior derecha de la cara frontal del cubo. En ella seguimos en el primer objetivo (antecedentes teóricos de la evaluación psicológica), pero ahora el nivel cognitivo es el tercero: solución de problemas (por ejemplo, el estudiante identifica errores en casos resueltos), y el nivel de complejidad de su modelo mental debe ser causal (debe explicar la interrelación funcional entre los elementos relevantes para la evaluación de casos de este dominio educativo).

Un Análisis Cognitivo de Tareas como el que se describe parte de la delimitación del dominio educativo, que debe implicar la identificación de las competencias deseables como desempeño en el dominio, la propuesta de indicadores de estas competencias, y posteriormente requiere de un trabajo fino de análisis en el cual se desglosan los objetivos (o competencias) del programa de estudio, y básicamente se identifican los niveles de complejidad: cognitiva, de tipo de conocimiento, del modelo mental a construir, etc., con un formato especial que se ha descrito por Castañeda (2006). Permite la especificación concreta de las habilidades que se evaluarán y se promoverán en la instrucción. Este análisis permite identificar cómo está conformado el conocimiento y, especialmente importante, cómo evaluar su estatus en términos del desempeño del aprendiz.

Con este esquema de evaluación se generan reactivos, actividades, simulaciones, sondeos, ejercicios y otras situaciones de evaluación interactiva en línea, sin perder de vista que cada elemento se puede ubicar dentro del esquema de análisis, y que al final se pueden realizar síntesis que arrojen una idea global del nivel de avance de cada aprendiz. La intención del uso de este esquema radica en que puede ser una alternativa de evaluación constante, que radiografe el proceso de aprendizaje.

En suma, el análisis permite identificar el nivel de complejidad de los componentes de las competencias. Evalúa conocimientos y habilidades a nivel micro, pero en el contexto de la competencia.

Castañeda (2004a) indica que la medición en sistemas orientados cognitivamente concibe los productos de aprendizaje como desarrollos graduales de habilidad cognitiva y plantea que el propósito de la medición es identificar el status actual del estudiante a partir de un modelo de desarrollo de la habilidad cognitiva. A partir del status y del tipo de errores detectados, será posible diseñar actividades de aprendizaje que fomenten el desarrollo cognitivo y los conocimientos pertinentes.

En el caso de construir instrumentos de evaluación con base en esta técnica, se sigue un procedimiento de validación y confiabilización, que consiste en lo siguiente:

1. el análisis inicial de competencias debe ser validado por expertos en el dominio, con algún instrumento y prueba de confiabilidad;
2. con un procedimiento similar se validan los niveles de demanda identificados para las tareas de las competencias;
3. con lo anterior se construyen los reactivos o ejercicios de evaluación, que una vez creados son sometidos también a la validación por jueces expertos, y se aplica una prueba estadística para determinar su grado de acuerdo;
4. cuando se cuenta con el instrumento, y antes de su utilización, se realiza un piloteo con una muestra de estudiantes, y a estos datos se aplican pruebas de calibración matemática, y
5. finalmente se depuran los instrumentos, que podrán convertirse a versiones Web e integrarse a ambientes de aprendizaje.

A partir de los elementos expuestos, en el presente trabajo se generó un conjunto de evaluaciones del desempeño de estudiantes inscritos en un curso en línea de psicología. El propósito del estudio fue poner a prueba una metodología para construir, validar y calibrar una batería de evaluaciones objetivas creadas con base en una técnica de Análisis Cognitivo de tareas.

2. MÉTODO

2.1. Participantes

Para el jueceo, los participantes fueron 10 profesores del área de Psicología Clínica de la carrera de psicología de la FES Iztacala, con una antigüedad promedio de 17 años, quienes han impartido al menos en diez ocasiones la asignatura de Psicología Clínica Teórica II, a la que corresponden los instrumentos desarrollados.

Para el piloteo y calibración de los instrumentos, los participantes fueron 70 alumnos de la licenciatura en psicología de la FESI-UNAM, inscritos en el sexto semestre de la carrera (43 mujeres y 27 hombres). Su edad promedio era de 22.3 años.

2.2. Situación y equipo

El jueceo se realizó en un cubículo de la Clínica Universitaria de Salud Iztacala, área en la que se concentra el cuerpo docente del área de Psicología Clínica. El cubículo tiene unas dimensiones de 3.5 por 2.5 metros, y cuenta con 1 computadora PC con procesador Pentium 4, 512 megabytes de memoria RAM, 120 Gigabytes de disco duro y acceso a Internet.

El piloteo de los instrumentos se realizó en una sala de cómputo ubicada en el Centro de Documentación de la Facultad de Estudios Superiores Iztacala. Las dimensiones de la sala son: 4 por 7 m² de superficie, paredes laterales de materiales prefabricados y un ventanal al frente. Esta sala cuenta con 40 computadoras PC con procesador Pentium 4, 256 megabytes de memoria RAM, 80 Gigabytes de disco duro y acceso a Internet.

2.3. Materiales

2.3.1. Construcción de los instrumentos

Para la construcción de los instrumentos, se utilizó la antología de la asignatura de Psicología Clínica Teórica II, desarrollada por un grupo de 12 profesores del área de Psicología Clínica de la carrera de Psicología de la FES Iztacala – UNAM.

La antología incluye dos volúmenes impresos: el primero es una guía didáctica, en la cual se incluyen: una introducción a la asignatura, así como introducciones a cada una de las tres unidades didácticas que la componen. En cada unidad se mencionan las referencias bibliográficas y se realiza una breve reseña de cada lectura. El segundo volumen es una recopilación impresa que integra reproducciones de todas las lecturas incluidas en el curso, agrupadas por unidad. Las unidades del curso son: 1) Introducción a la aproximación cognitiva conductual, con dos lecturas; 2) La evaluación cognitivo-conductual, con tres lecturas, y 3) El análisis funcional, con tres lecturas.

Se requirieron las siguientes herramientas para la creación de la versión en línea de los instrumentos de evaluación del curso:

Lenguaje HTML, para hacer las páginas Web que se requirieron para publicar el instrumento.

Base de datos MySQL, que permitió almacenar las respuestas a los reactivos por parte de cada uno de los participantes; lenguaje Java, con el objeto de programar los cuestionarios para el ambiente Web, y la

comunicación con la base de datos para almacenar las respuestas y posteriormente realizar reportes. Un servidor conectado a Internet en el cual hospedar el instrumento EDAOM en línea. Esto se hizo con el proveedor de almacenamiento para Internet www.performancehosting.net. Un software FTP (File Transfer Protocol) para transferir los archivos de la aplicación en el servidor del sitio Web: CuteFTP de GlobalScape.

FIGURA 1. MENÚ DE INGRESO A LAS EVALUACIONES DESDE EL META-TUTOR

Actividades de Samantha Arceo Rodríguez	
Asignatura 2	Psicología Clínica Teórica I
Evaluación 1	Evaluación inicial de opiniones acerca del estudio.
Evaluación 2	Evaluación inicial de conocimientos (no tiene ningún peso en la calificación).
Evaluación 3	Evaluación del tema: La aproximación Cognitiva Conductual.
Evaluación 4	Evaluación del tema: Evaluación y entrevista.
Evaluación 5	Evaluación del tema: Evaluación y Análisis funcional.
Tema 1	La aproximación cognitiva conductual
Tema 2	Evaluación y entrevista
Tema 3	Evaluación y análisis funcional

Meta-Tutor. Los instrumentos de evaluación en línea se incluyó como un componente dentro del ambiente de aprendizaje *Meta-Tutor* (Peñalosa y Castañeda, 2008b), que brindaba otros servicios instruccionales, pero al inicio presentaba un menú donde, entre otras opciones, permitía ingresar al EDAOM en línea cuando se programaba su aplicación.

FIGURA 2. MUESTRA DE UNO DE LOS INSTRUMENTOS DE EVALUACIÓN EN LÍNEA

Actividades del usuario - Microsoft Internet Explorer

http://www.metatutor.com/evaluacion.jsp

meta.tutor Ambiente de aprendizaje

Evaluación de Samantha Arceo Rodríguez

6 De acuerdo a la estructura de antecedentes del diagrama, elige la opción que engloba los factores a los que da lugar de manera directa la situación ilustrada en el siguiente texto:

"La psicoterapia americana estuvo muy influida por la psicología experimental. Empezando por John Watson e Ivan Pavlov, y continuando por Thorndike y Skinner, el centro de estudio se desplazó de la introspección y el estudio de la mente, hacia el desarrollo y examen de los principios del aprendizaje y el cambio de la conducta. En este proceso, la mente se redujo a una condición de epifenómeno de la "caja negra", no porque no fuera importante o porque fuera inexistente, sino porque no podía observarse por otros individuos (Skinner, 1953). El conductismo radical subraya así dos principios explicativos: el condicionamiento clásico, basado en el aprendizaje por asociación, y el condicionamiento operante, basado en el aprendizaje por las consecuencias. La terapia conductual desarrolló intervenciones desde ambos principios y se convirtió en una fuerza importante en la psicología clínica americana durante 30 años. Sin embargo, diversos factores fueron responsables de su paulatina pérdida de importancia. En primer lugar, aunque las técnicas conductistas demostraban su eficacia, no se hacían por encima de otras alternativas terapéuticas. Precisamente porque los psicólogos conductistas hicieron investigaciones empíricas, se desconcertaron cuando sus resultados mostraron diferencias no significativas con otras intervenciones. En segundo lugar, las técnicas conductistas no siempre funcionaban -contrariamente a las predicciones derivadas de las leyes del aprendizaje-. En otras palabras, la "caja negra" entre los estímulos y las respuestas y sus consecuencias parecía ser la responsable de mucha más variación de lo que podía pensarse! Por tanto, empezó la búsqueda de una explicación para esa variabilidad." (Dowd, 1997, pp. 26-27).

Antecedentes de la Aproximación Cognitiva Conductual

Se dividen en:

- 1 Histórico-Filosóficos
- 2 Factores dentro de la psicología que propician su aparición
- 3 Surgimiento de las terapias cognitivas

☐ 1 Factores histórico-filosóficos

☐ 2: desarrollo de factores dentro de la psicología que propician la aparición de la TCC

☒ 3: surgimiento de las terapias cognitivas

Siguiente

La versión final de los instrumentos de evaluación de conocimiento académico aparecía en el *Meta-Tutor*, podía accederse a los mismos a través de un índice como el que se muestra en la figura 1, y un ejemplo de un reactivo de un instrumento aparece en la figura 2.

2.3.2. Jueceo de los instrumentos

Para el jueceo de los instrumentos, se requirió por un lado de los instrumentos ya desarrollados y disponibles en la dirección de Internet del *Meta-Tutor*. Se utilizó el formato de jueceo que se muestra en la tabla 1.

En el eje horizontal aparecen los números de los reactivos, y en el vertical las claves de cada uno de los jueces. Cada casilla es un espacio para la calificación del reactivo (cero o uno, dependiente del acuerdo o desacuerdo del juez con la aseveración que se evalúa).

TABLA1. FORMATO PARA JUECEO DE INSTRUMENTOS

	i1	i2	i3	i4	i5	i6	i7	i8	i9	i10	i11	i12	i13	i14	i15	i16	i17	i18	i19	i20	i21	i22	i23	i24	i25	i26	i27
j1																											
j2																											
j3																											
j4																											
j5																											
j6																											
j7																											
j8																											
j9																											
j10																											

2.3.3. Calibración de reactivos

Para la calibración de los instrumentos, se utilizaron dos programas para PC: 1) ITEMAN (1993), un software para análisis de reactivos desarrollado de acuerdo con la teoría clásica de tests, de la casa *Assessment Systems Corporation*, y 2) RASCAL (1992), un software para calibración de reactivos de acuerdo con el modelo *Rasch* de un parámetro, de la Teoría de Respuesta al Ítem, de la misma casa de desarrollo de software.

2.4. Procedimiento

2.4.1. Construcción de los instrumentos

Los instrumentos que se construyeron fueron:

1. pre-postest, un cuestionario de 27 preguntas relacionadas con las tres unidades didácticas del curso de Psicología Clínica Teórica II (PCTII).
2. Evaluación de la unidad I, un cuestionario de 48 preguntas relacionadas con la primera unidad didáctica del curso de PCTII: la aproximación cognitivo conductual.
3. Evaluación de la *unidad 2*, un cuestionario de 40 preguntas relacionadas con la segunda unidad didáctica del curso de PCTII: la evaluación cognitivo conductual.

4. Evaluación de la *unidad 2*, un cuestionario de 38 preguntas relacionadas con la tercera unidad didáctica del curso de PCTII: el análisis funcional.

Los instrumentos enumerados se construyeron siguiendo la lógica que se expone a continuación.

2.4.2. Prueba Pre-Post

Evaluaba en general el conocimiento del dominio, permitió conocer el nivel inicial de los participantes, y compararlo con su nivel final.

Pasos realizados en el desarrollo de la evaluación pre – post.

1. Se realizó un análisis cognitivo de tareas como se describe en el apartado homónimo en este trabajo, a través del cual se identificaron los componentes de tarea a ser evaluados. Con ellos se construyó el diseño que dio el sustento teórico y puso los límites al universo de contenido que sería evaluado. Como productos derivados del diseño de observación, se establecieron los tres modelos mentales y los tres niveles de abstracción que representarían los contenidos cubiertos por la asignatura Psicología Clínica Teórica II de la carrera de psicología de la FES Iztacala, elegida como el área de interés.
2. El diseño de observación representa los parámetros que rigieron la construcción de la evaluación. Éste se muestra en la tabla 2.

TABLA 2. ESPACIO DE CONTENIDOS DE LA EVALUACIÓN PRE - POST

		Modelo mental		
		Conceptual	Estructural	Causal
Niveles de Complejidad cognitiva	Solución de problemas	1	4	7
	Aplicación del conocimiento	2	5	8
	Comprensión	3	6	9

El dominio de conocimiento de la asignatura a impartir en el sistema de aprendizaje en línea se dividió en tres cortes, de acuerdo con lo expuesto en la sección que describe el ACT. Se generaron reactivos representativos del conocimiento para cada casilla en cada unidad. Son 3 unidades, por lo tanto para esta evaluación en total se generaron 27 reactivos.

2.4.3. Evaluaciones de las unidades 1, 2 y 3

La evaluación pre-post se desarrolló tomando en cuenta como ejes del diseño de observación tres modelos mentales y tres niveles de abstracción, todo lo anterior de tres unidades. En las evaluaciones de las unidades 1, 2 y 3, el diseño de observación tuvo una ligera diferencia: el universo del dominio se limitaba a los objetivos de cada una de las unidades, pero también tomando en cuenta los tres modelos mentales y los niveles de complejidad cognitiva.

Los tres instrumentos enumerados se generaron con la misma metodología, con base en la descripción del ACT, como se ilustra en la figura 1.

Como se ha mencionado, en esta figura se observan los gradientes de complejidad cognitiva y de modelos mentales, pero el tercer eje del modelo tridimensional se relaciona con la complejidad de contenidos de las unidades 1, 2 y 3 del curso. A continuación se describen los contenidos de estas unidades, así como los reactivos construidos para cada una.

Objetivos de las unidades:

Las unidades de aprendizaje tenían los siguientes objetivos, de acuerdo con la guía didáctica y antología del curso:

Unidad 1. La aproximación cognitivo conductual.

Objetivos. Al final del curso, el alumno:

- Conocerá los aspectos históricos de la Aproximación Cognitiva Conductual (ACC)
- Conocerá la definición y explicará los supuestos de la ACC
- Comprenderá las terapias y técnicas disponibles, sus usos y características

Se crearon dos reactivos por cada nivel de complejidad cognitiva, con dos niveles de dificultad en cada uno, y esto para cada uno de los modelos mentales y objetivos de la unidad excepto en los casos en los que no aplicara el cruce, de manera que se generaron 48 reactivos.

Unidad 2. La evaluación cognitivo - conductual

Las unidades de aprendizaje tenían los siguientes objetivos, de acuerdo con la guía didáctica y antología del curso:

- Caracterización de la Evaluación Conductual
- Instrumentos de evaluación de la TCC
- Conceptualización de la Entrevista conductual

Se crearon dos reactivos por cada nivel de complejidad cognitiva, con dos niveles de dificultad en cada uno, y esto para cada uno de los modelos mentales y objetivos de la unidad excepto en los casos en los que no aplicara el cruce, de manera que se crearon 40 reactivos.

Unidad 3. Evaluación y análisis funcional

Las unidades de aprendizaje tenían los siguientes objetivos, de acuerdo con la guía didáctica y antología del curso:

- Pasos en el desarrollo de la evaluación conductual
- Análisis funcional de casos

Se crearon dos reactivos por cada nivel de complejidad cognitiva, con dos niveles de dificultad en cada uno, y esto para cada uno de los modelos mentales y objetivos de la unidad. En este caso, sólo eran dos unidades, por lo tanto se generaron 36 reactivos.

2.4.4. Procedimiento de jueceo de los instrumentos

El procedimiento de jueceo fue el mismo en los cuatro casos, como sigue:

Se realizó la validación de los cuatro instrumentos de evaluación mediante la revisión por parte de un grupo de 10 jueces expertos en la asignatura, que son profesores con una experiencia de más de 10 años impartiendo la misma.

Se entregó el instrumento a los jueces, y se les dio un formato en el que debían indicar con un "1" si estaban de acuerdo con la validez de la pregunta, y un "0" en caso contrario.

Los jueces tuvieron acceso al diseño de observación, previamente elaborado, y en función de él realizaron sus observaciones acerca de la validez. Además de sus calificaciones de cada reactivo, todos los comentarios de los jueces se anotaron para efectos de optimización de la calidad de los instrumentos.

La prueba de confiabilidad de Kuder-Richardson-20, ilustrada en la Ecuación 2.1, que permite evaluar homogeneidad o consistencia interna en pruebas con datos dicotómicos, fue aplicada a los resultados del jueceo de los cuatro instrumentos, con el objeto de determinar el grado de acuerdo entre los jueces.

$$r_{kk} = \frac{k}{k-1} \left(\frac{s_x^2 - \sum p_i q_i}{s_x^2} \right) \quad (1)$$

En donde:

k= número de reactivos de la prueba

s_x^2 = varianza de los resultados de la prueba

p_i = proporción de jueces que aceptan un reactivo

$q_i = 1 - p_i$ = proporción de quienes no aceptan el reactivo

La prueba de Kuder-Richardson 20 entrega datos de la homogeneidad de la prueba; esto es, la consistencia interna de la misma (Brown, 1980); por lo tanto, serán útiles en este estudio para valorar si las opiniones de los jueces son consistentes.

2.4.5. Procedimiento de calibración de reactivos

La calibración es "un procedimiento psicométrico que permite poner en su punto de funcionamiento óptimo a los ítems de un instrumento, con base en un modelo de medición, de manera que al aplicarse a una población meta, aporten con adecuación, confiabilidad y utilidad, la información necesaria para tomar decisiones o medidas que, por ejemplo, conduzcan a la atención de un problema o a la formulación de un diagnóstico" (López, 2006a: 43).

El procedimiento general de calibración utilizado consistió en los siguientes pasos:

1. Al concluir el estudio, descargar del sitio Web del *Meta-Tutor* la información correspondiente a la lista de respuestas de 70 de los participantes ante cada uno de los instrumentos a calibrar.
2. Generar el archivo de base de datos para análisis, de acuerdo con los requerimientos de los programas para calibración descritos. Todos los programas requieren de una configuración de los datos para su lectura en archivos '.prn', con la distribución que se muestra en la tabla 3, donde las líneas: 1) primera línea, llamada de control, estipula: el número de reactivos, códigos para reactivos omitidos, no concluidos, así como la clave de usuarios; 2) en la segunda línea aparecen las respuestas correctas del instrumento; 3) en la tercera línea el número de opciones de respuesta, 4) en la cuarta línea se indica si se incluirá o no el reactivo en el análisis, y 5) de la quinta línea en adelante, aparecen todas las respuestas emitidas a cada reactivo por cada uno de los participantes.

TABLA 3. FORMATO DE ENTRADA DE DATOS PARA CALIBRACIÓN

27 0 N 0
233323123232122231321131122
3333333333333333333333333333
YYYYYYYYYYYYYYYYYYYYYYYYYYY
311121233112132213332131123
311332112233122322233322131
212331133212131221223323133
11112111213212323232232311
112321233322112233223312223
313331112232112223321333133
311332212232111231122111221
211312122112121233332112131

1. Ejecutar el programa en cuestión (dependiendo del análisis que se esté realizando), ITEMAN o RASCAL, indicando el nombre del archivo con los datos similar al descrito en la tabla 3, así como la ruta donde éste se encuentre en la computadora donde se ejecute el análisis.
2. El programa entrega un archivo con extensión '.out', que incluye la valoración de cada uno de los reactivos calibrados, de acuerdo con los parámetros que cada software maneja.

Los análisis realizados con los programas de calibración de reactivos enumerados anteriormente se describen a continuación.

ITEMAN y la Teoría Clásica de los Tests, o Teoría de Respuesta a la Prueba (TRP)

La Teoría Clásica de los Tests fue utilizada originalmente para desarrollar pruebas que determinaran la habilidad de las personas con respecto a ciertos reactivos incluidos en la prueba que pudieran satisfacer ciertas características y criterios predeterminados. En la TRP, las estadísticas de los reactivos dependen de las habilidades de los participantes; en otras palabras, los estudiantes con alta habilidad hacen que los reactivos se vean fáciles, y los alumnos con baja habilidad hacen que parezcan difíciles. Los índices de discriminabilidad se relacionan directamente con la homogeneidad del grupo de estudiantes evaluados.

ITEMAN Es un programa especializado en el análisis de reactivos de acuerdo con la Teoría de la Respuesta a la Prueba. López (2006b) describe a este tipo de programas indicando que ofrecen información útil acerca de los niveles de dificultad de cada reactivo, la discriminación entre desempeños altos y bajos, la distribución de respuestas en las opciones o distractores de cada reactivo, la correlación entre la respuesta al reactivo con respecto a las respuestas a toda la prueba, así como un sumario estadístico con los valores principales del instrumento aplicado, como el número de reactivos, examinados, media, desviación estándar, varianza, datos de la distribución, etc.

2.4.6. Programas de calibración según la Teoría de Respuesta al ÍTEM (TRI)

La TRI postula las siguientes seis ideas básicas: a) el desempeño de un evaluado en un reactivo de la prueba puede ser explicado por un conjunto de factores llamados rasgos, rasgos latentes o habilidades; b) los parámetros de los reactivos son independientes de la población de evaluados; en otras palabras, las estimaciones de la habilidad del evaluado no dependen de la prueba ni del grupo; c) los parámetros de habilidad son independientes de la elección de reactivos, los parámetros de reactivos y habilidades son independientes; d) la precisión de la estimación de habilidad puede determinarse en cada nivel de

habilidad; e) pueden seleccionarse reactivos para conformar una prueba con ciertas características, y f) la relación entre el desempeño de un participante en un reactivo y el conjunto de rasgos que subyacen al desempeño ante el reactivo pueden describirse mediante una función llamada función o curva característica del reactivo (Hambleton, Rogers y Swaminathan, 1991).

En la TRI, un reactivo puede caracterizarse por tres parámetros: a) el valor de discriminación, b) el nivel de dificultad, y c) el nivel de adivinación.

López (2006b) indica que estos modelos psicométricos permiten estimar el nivel de la habilidad de un examinando sobre el rasgo que está siendo medido, independientemente de la prueba específica que responda, asegurando que lo que se mide no dependa de con qué o a quién se mida.

El programa RASCAL se basa en el modelo Rasch de la TRI, cuyos supuestos son planteados por López (2006b):

1. Que la habilidad de los sustentantes puede determinarse de manera confiable, sin importar los reactivos específicos que se utilicen en una prueba en particular, por lo que se rompe la dependencia hacia el instrumento particular empleado para hacer la medición siempre y cuando los ítems hayan pasado por un proceso cuidadoso para su calibración. Este problema señala una de las mayores debilidades de la Teoría Clásica de Respuesta a la Prueba.
2. Que los reactivos calibrados tendrán también un comportamiento muy estable en sus valores de dificultad, independientemente de los sujetos particulares a los que se aplique el instrumento, en poblaciones que sean razonablemente similares.

El parámetro fundamental de análisis del programa RASCAL es la dificultad del reactivo, que implica el nivel de habilidad del participante. En términos generales, el rango de dificultad está entre -2 y 2, y los valores más allá de este rango representarán reactivos muy fáciles o muy difíciles. Se dice que esta escala permite identificar el ajuste a la prueba por parte de los evaluados, así como su habilidad.

El programa RASCAL permite también generar una curva característica de la prueba, en la que podrían identificarse los niveles de habilidad y la posibilidad de acertar.

El procedimiento de calibración incluyó el uso de los dos programas mencionados (ITEMAN y RASCAL) para el análisis de calibración de reactivos. Tomando en cuenta la propuesta de López (2006b) en el sentido de integrar en una sola tabla los indicadores de ambos modelos, en el presente estudio se integra información de calibración de los modelos de la TRP, así como de la TRI, con el objeto de tomar decisiones acerca de la constitución de los instrumentos.

TABLA 4. FORMATO DE CALIBRACIÓN CON INFORMACIÓN DE TRP Y TRI

Item	1	2	3	4	5	6	7	8	9	10
Dificultad	0.46	0.14	0.42	0.72	0.54	0.21	0.70	0.38	0.25	0.48
Discriminación	0.02	0.19	0.36	0.22	0.33	0.33	0.09	0.48	-0.06	0.36
Correlación	-0.03	0.16	0.25	0.29	0.30	0.40	0.01	0.38	0.06	0.27
Dific. TRI	-0.03	1.63	0.14	-1.11	-0.32	1.14	-1.05	0.32	0.91	-0.09
Error Estándar	0.24	0.34	0.25	0.27	0.24	0.29	0.26	0.25	0.28	0.24

La tabla 4 es un ejemplo de cómo se reportan los resultados de las calibraciones. En la primera fila se incluye el número de reactivo, en las filas 2 a 4 la información derivada de los parámetros de la TRP y el software ITEMAN, y en las filas 5 y 6 la información de la TRI, obtenida con el software RASCAL. Con la

información de la tabla será posible identificar las propiedades de cada reactivo de una manera integrada, y en todo caso tomar las decisiones de la pertinencia o eliminación de reactivos en cada uno de los instrumentos a analizar.

3. RESULTADOS DEL JUECEO Y ANÁLISIS DE REACTIVOS

Toda evaluación enfrenta desafíos técnicos y tensiones que debe resolver. Para el caso de educación de adultos hay dificultades específicas relacionadas con una escasa cultura evaluativa, con nociones y prácticas erradas en evaluación, con la heterogeneidad y complejidad de la población que participa en las ofertas educativas, con la falta de formación docente y escasa reflexión pedagógica acerca de la evaluación.

3.1. Instrumento pre-post

3.1.1. Jueceo

La tabla 5 muestra las valoraciones de los jueces (j1.. j10) con respecto a los reactivos (i1 – i27) del instrumento pre-post:

TABLA 5. DATOS DE JUECEO DEL INSTRUMENTO PRE - POST

	i1	i2	i3	i4	i5	i6	i7	i8	i9	i10	i11	i12	i13	i14	i15	i16	i17	i18	i19	i20	i21	i22	i23	i24	i25	i26	i27
j9	0	1	1	1	1	1	1	0	1	1	1	1	0	1	1	0	1	1	1	0	1	1	1	1	1	1	1
j10	1	0	1	1	1	1	0	1	0	1	1	1	1	0	1	0	1	1	1	1	0	1	1	0	1	1	0
j3	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1
j7	1	1	0	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
j8	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
j1	1	1	1	1	1	1	1	1	0	1	1	0	1	1	1	1	1	0	1	1	1	0	1	1	1	1	1
j2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
j6	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
j5	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
j4	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1

La tabla 6 muestra los resultados de las diferentes medidas utilizadas en la prueba Kuder-Richardson 20.

TABLA 6. VALORES DE LA PRUEBA KUDER-RICHARDSON 20 EN PRE-POSTEST

k	27
$\sum p q_i$	1.74
$\sum p_i^2$	25
s_x^2	7
r_{kk}	.83

Se observa que la prueba arrojó un coeficiente de confiabilidad de .83, que representa un valor aceptable de acuerdo entre jueces, e indica que no existen diferencias significativas entre sus opiniones respecto a la pertinencia de los reactivos.

TABLA 7. ANÁLISIS DE REACTIVOS DEL INSTRUMENTO PRETEST

Item	1	2	3	4	5	6	7	8	9	10	11	12	13
Dificultad	0.46	0.14	0.42	0.72	0.54	0.21	0.70	0.38	0.25	0.48	0.46	0.70	0.59
Discriminación	0.02	0.19	0.36	0.22	0.33	0.33	0.09	0.48	-0.06	0.36	0.19	0.21	0.16
Correlación	-0.03	0.16	0.25	0.29	0.30	0.40	0.01	0.38	0.06	0.27	0.20	0.31	0.11
<i>Dific. TRI</i>	<i>-0.03</i>	<i>1.63</i>	<i>0.14</i>	<i>-1.11</i>	<i>-0.32</i>	<i>1.14</i>	<i>-1.05</i>	<i>0.32</i>	<i>0.91</i>	<i>-0.09</i>	<i>-0.03</i>	<i>-1.05</i>	<i>-0.55</i>
<i>Error Estándar</i>	<i>0.24</i>	<i>0.34</i>	<i>0.25</i>	<i>0.27</i>	<i>0.24</i>	<i>0.29</i>	<i>0.26</i>	<i>0.25</i>	<i>0.28</i>	<i>0.24</i>	<i>0.24</i>	<i>0.26</i>	<i>0.25</i>

Item	14	15	16	17	18	19	20	21	22	23	24	25	26	27
Dificultad	0.42	0.46	0.79	0.34	0.30	0.56	0.42	0.28	0.37	0.39	0.34	0.87	0.45	0.28
Discriminación	-0.13	0.14	0.06	0.12	0.29	0.35	0.01	0.13	0.12	0.48	0.34	0.07	0.15	0.12
Correlación	0.05	0.15	0.08	0.04	0.24	0.34	-0.01	0.13	0.21	0.35	0.38	0.07	0.22	0.16
<i>Dific. TRI</i>	<i>0.14</i>	<i>-0.03</i>	<i>-1.49</i>	<i>0.50</i>	<i>0.70</i>	<i>-0.43</i>	<i>0.14</i>	<i>0.76</i>	<i>0.38</i>	<i>0.26</i>	<i>0.50</i>	<i>-2.10</i>	<i>0.02</i>	<i>0.76</i>
<i>Error Estándar</i>	<i>0.25</i>	<i>0.24</i>	<i>0.29</i>	<i>0.26</i>	<i>0.26</i>	<i>0.24</i>	<i>0.25</i>	<i>0.27</i>	<i>0.25</i>	<i>0.25</i>	<i>0.26</i>	<i>0.35</i>	<i>0.24</i>	<i>0.27</i>

Análisis de reactivos

Como se indicó en la sección de procedimiento, el análisis de reactivos del instrumento de evaluación pre-post fue realizado con ayuda del programa ITEMAN, así como el RASCAL, arrojando en resumen los datos que se observan en la tabla 7. Las filas en cursivas representan los valores del RASCAL, y las filas que no están en cursivas son valores del ITEMAN.

3.1.2. Ajustes al instrumento después del análisis

Debido a que esta aplicación de este instrumento corresponde a un conjunto de participantes con un estatus de principiantes, la dificultad identificada por ITEMAN es alta, con una media de .46; la discriminación es baja en promedio (.19), así como la correlación promedio de cada reactivo con el resto (.19). Por otro lado, la dificultad de acuerdo con la TRI está en un rango de -2.10 hasta 1.63, lo cual ubica a este rango como apto para identificar el ajuste de los evaluados.

Dado que un criterio para seleccionar reactivos es que la discriminación no sea negativa, se decidió eliminar los reactivos 9 y 14, que son negativos en este parámetro.

Por lo tanto, para este instrumento se dejan 25 reactivos en total.

3.2. Instrumento de evaluación de la Unidad I

Los resultados del jueceo del instrumento de evaluación de la Unidad 1 se muestran en la tabla 8, que se muestra abajo.

TABLA 8. DATOS DE JUECEO DEL INSTRUMENTO DE LA UNIDAD 1

	i1	i2	i3	i4	i5	i6	i7	i8	i9	i10	i11	i12	i13	i14	i15	i16	i17	i18	i19	i20	i21	i22	i23	i24
j1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
j2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
j3	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	0
j4	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
j5	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
j6	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
j7	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
j8	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
j9	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1

j10	0	1	1	0	1	1	0	1	0	1	0	1	1	0	1	1	0	1	1	0	1	1	0	1
	i25	i26	i27	i28	i29	i30	i31	i32	i33	i34	i35	i36	i37	i38	i39	i40	i41	i42	i43	i44	i45	i46	i47	i48
j1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1
j2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
j3	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1
j4	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
j5	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
j6	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
j7	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
j8	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
j9	1	1	1	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
j10	0	1	1	0	1	1	0	1	1	1	1	1	1	1	0	1	1	1	0	1	1	1	0	0

La tabla 9 muestra los resultados de las diferentes medidas utilizadas en la prueba Kuder-Richardson 20.

TABLA 9. VALORES DE LA PRUEBA KUDER-RICHARDSON EN UNIDAD I

k	48
$\sum p_i q_i$	1.13
$\sum p_i^2$	42.40
sx ²	21.86
rkk	.97

Se observa que la prueba arrojó un coeficiente de confiabilidad de .97, que representa un grado importante de acuerdo entre jueces, e indica que no existen diferencias significativas entre sus opiniones respecto a la pertinencia de los reactivos.

3.2.1. Análisis de reactivos instrumento Unidad 1

Como se ha dicho antes, el instrumento de la Unidad 1 contaba con 48 reactivos, mismos que fueron sometidos al análisis ya descrito, donde se realizaron los procedimientos de calibración de acuerdo con la TRP (ITEMAN) y la TRI (RASCAL). Los resultados se muestran en la tabla 10.

En el análisis de acuerdo con el modelo de la TRP, realizado con el programa ITEMAN, La mayoría de los reactivos demostraron valores adecuados en función de los parámetros utilizados. Sin embargo, los reactivos 6, 27, 40 y 46 tuvieron una correlación negativa; los reactivos 27, 37, 40 y 46 tuvieron una discriminación negativa.

Por otro lado, al aplicar el análisis de acuerdo con el modelo de TRI, realizado con el programa RASCAL, los índices de discriminación fueron adecuados, excepto en los casos de los reactivos 6, 25, 32 y 46, con valores extremos de 4.01, -2.39, -2.39 y 2.32, respectivamente.

TABLA 10. ANÁLISIS DE REACTIVOS. INSTRUMENTO DE LA UNIDAD I

Item	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Dif.	0.27	0.85	0.52	0.44	0.88	0.05	0.64	0.79	0.56	0.79	0.50	0.80	0.79	0.74	0.58	0.68
Disc.	0.09	0.35	0.26	0.27	0.25	0.00	0.17	0.25	0.12	0.25	0.27	0.20	0.25	0.41	0.26	0.55
Corr.	0.22	0.39	0.20	0.26	0.35	-0.04	0.20	0.29	0.09	0.21	0.28	0.22	0.22	0.35	0.30	0.46
Dific. TRI	1.87	-1.03	0.75	1.08	-1.30	4.01	0.21	-0.60	0.55	-0.60	0.82	-0.69	-0.60	-0.33	0.49	-0.01
Error Est.	0.29	0.35	0.26	0.26	0.39	0.60	0.27	0.31	0.26	0.31	0.26	0.32	0.31	0.29	0.26	0.27

Item	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32
Dif.	0.77	0.32	0.85	0.73	0.88	0.70	0.24	0.73	0.95	0.91	0.29	0.33	0.82	0.41	0.92	0.95
Disc.	0.45	0.28	0.15	0.26	0.15	0.45	0.24	0.50	0.15	0.05	-0.06	0.08	0.10	0.52	0.15	0.10
Corr.	0.49	0.30	0.33	0.34	0.15	0.52	0.18	0.42	0.33	0.03	-0.01	0.14	0.07	0.38	0.25	0.19
Dific. TRI	-0.50	1.64	-1.03	-0.24	-1.30	-0.08	2.04	-0.24	-2.39	-1.63	1.79	1.56	-0.80	1.21	-1.84	-2.39
Error Est.	0.30	0.27	0.35	0.29	0.39	0.28	0.30	0.29	0.60	0.44	0.28	0.27	0.33	0.26	0.47	0.60

Item	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48
Dif.	0.83	0.38	0.76	0.95	0.77	0.68	0.39	0.58	0.89	0.58	0.74	0.88	0.55	0.20	0.82	0.42
Disc.	0.25	0.41	0.01	0.15	-0.04	0.41	0.27	-0.12	0.10	0.36	0.40	0.10	0.17	-0.01	0.30	0.17
Corr.	0.40	0.26	0.00	0.20	0.10	0.39	0.26	-0.01	0.24	0.28	0.40	0.16	0.19	-0.06	0.27	0.12
Dific. TRI	-0.91	1.35	-0.41	-2.39	-0.50	-0.01	1.28	0.49	-1.46	0.49	-0.33	-1.30	0.62	2.32	-0.80	1.15
Error Est.	0.34	0.26	0.30	0.60	0.30	0.27	0.26	0.26	0.41	0.26	0.29	0.39	0.26	0.32	0.33	0.26

3.2.2. Ajustes al instrumento de la Unidad 1 después del análisis

Dados los datos indicados anteriormente, se decidió eliminar los siguientes reactivos: 6, 25, 27, 32, 36, 37, 40 y 46.

Por lo tanto, este instrumento se conformó de 40 reactivos.

3.3. Instrumento de evaluación de la Unidad 2

Los resultados del jueceo del instrumento de evaluación de la Unidad 2 se muestran en la tabla 11.

TABLA 11. DATOS DE JUECO DEL INSTRUMENTO DE LA UNIDAD 2

	i1	i2	i3	i4	i5	i6	i7	i8	i9	i10	i11	i12	i13	i14	i15	i16	i17	i18	i19	i20
j1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
j2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
j3	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1
j4	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1
j5	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
j6	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
j7	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
j8	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1
j9	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
j10	1	1	1	1	1	1	0	1	1	0	1	1	1	1	0	1	1	1	0	1

	i21	i22	i23	i24	i25	i26	i27	i28	i29	i30	i31	i32	i33	i34	i35	i36	i37	i38	i39	i40
j1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	0	0
j2	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1
j3	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
j4	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
j5	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1
j6	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
j7	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1
j8	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
j9	1	1	1	1	1	1	1	1	1	1	0	1	1	1	0	1	0	1	1	1
j10	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	0	1	1

La tabla 12 muestra los resultados de las diferentes medidas utilizadas en la prueba Kuder-Richardson 20.

TABLA 12. VALORES DE LA PRUEBA KUDER-RICHARDSON EN UNIDAD 2

k	40
\square piqi	0.73
\square	34.8
sx2	14.26
rkk	0.973

Se observa que la prueba arrojó un coeficiente de confiabilidad de .97, que representa un grado importante de acuerdo entre jueces, e indica que no existen diferencias significativas entre sus opiniones respecto a la pertinencia de los reactivos.

3.3.1. Análisis de reactivos del instrumento Unidad 2

El instrumento de la *Unidad 2* contaba con 40 reactivos, mismos que fueron sometidos al análisis de calibración de acuerdo con la TRP (ITEMAN) y la TRI (RASCAL). Los resultados se muestran en la tabla 13.

TABLA 13. ANÁLISIS DE REACTIVOS DEL INSTRUMENTO DE LA UNIDAD 2

Item	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Dif.	0.85	0.80	0.24	0.48	0.98	0.79	0.76	0.58	0.89	0.79	0.70	0.56	0.50	0.21
Disc.	0.17	0.41	-0.05	0.52	0.05	0.35	0.30	0.72	0.23	0.17	0.28	0.16	0.33	-0.30
Corr.	0.16	0.44	0.12	0.42	0.26	0.22	0.28	0.57	0.37	0.22	0.24	0.12	0.26	-0.22
Dific. TRI	-1.18	-0.85	1.89	0.72	-3.68	-0.75	-0.57	0.33	-1.61	-0.75	-0.24	0.39	0.66	2.08
Error Est.	0.35	0.32	0.30	0.26	1.01	0.31	0.30	0.26	0.41	0.31	0.28	0.26	0.26	0.31

Item	15	16	17	18	19	20	21	22	23	24	25	26	27	28
Dif.	0.62	0.62	0.77	1.00	0.41	0.59	0.62	0.39	0.55	0.65	0.89	0.44	0.52	0.42
Disc.	0.44	0.52	0.16	0.00	0.32	0.28	0.03	0.44	0.11	0.32	0.18	0.48	0.42	0.66
Corr.	0.30	0.42	0.11	0.00	0.31	0.23	0.10	0.39	0.16	0.33	0.24	0.41	0.34	0.57
Dific. TRI	0.12	0.12	-0.66	borrar	1.06	0.26	0.12	1.12	0.46	-0.02	-1.61	0.92	0.59	0.99
Error Est.	0.26	0.26	0.30		0.26	0.26	0.26	0.26	0.26	0.27	0.41	0.26	0.26	0.26

Item	29	30	31	32	33	34	35	36	37	38	39	40
Dif.	0.35	0.50	0.77	0.47	0.91	0.62	0.67	0.67	0.76	0.76	0.65	0.42
Disc.	0.20	0.31	0.22	0.18	0.17	0.29	0.50	-0.04	0.15	0.21	-0.10	0.40
Corr.	0.14	0.20	0.14	0.20	0.25	0.25	0.45	0.01	0.17	0.21	0.01	0.42
Dific. TRI	1.34	0.66	-0.66	0.79	-1.78	0.12	-0.09	-0.09	-0.57	-0.57	-0.02	0.99
Error Est.	0.27	0.26	0.30	0.26	0.43	0.26	0.27	0.27	0.30	0.30	0.27	0.26

En el análisis de acuerdo con el modelo de la TRP, realizado con el programa ITEMAN, el reactivo 14 tuvo una correlación negativa con el resto de los reactivos; los reactivos 3, 14, 36 y 39 tuvieron una discriminación negativa. El reactivo 18 tuvo una dificultad de 1, ya que todos los participantes de la muestra lo respondieron correctamente.

Por otro lado, al aplicar el análisis de acuerdo con el modelo de TRI, realizado con el programa RASCAL, los índices de discriminación fueron adecuados, excepto en el reactivo 14, con un valor extremo de 2.08, lo que lo hacía demasiado difícil.

3.3.2. Ajustes al instrumento de la Unidad 1 después del análisis

Dados los datos indicados anteriormente, se decidió eliminar los siguientes reactivos: 3, 5, 14, 18, 36 y 39.

Por lo tanto, este instrumento se conformó de 34 reactivos.

3.4. Instrumento de evaluación de la Unidad 3

Los resultados del jueceo del instrumento de evaluación de la Unidad 3 se muestran en la tabla 14.

TABLA 14. DATOS DE JUECO DEL INSTRUMENTO DE LA UNIDAD 3

	i1	i2	i3	i4	i5	i6	i7	i8	i9	i10	i11	i12	i13	i14	i15	i16	i17	i18	i19
j1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
j2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
j3	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
j4	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
j5	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
j6	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
j7	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
j8	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
j9	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
j10	1	1	1	1	0	1	1	1	1	0	1	1	1	1	1	1	0	1	1

	i20	i21	i22	i23	i24	i25	i26	i27	i28	i29	i30	i31	i32	i33	i34	i35	i36
j1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1
j2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
j3	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
j4	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
j5	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
j6	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
j7	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
j8	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1
j9	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1
j10	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1

La tabla 15 muestra los resultados de las diferentes medidas utilizadas en la prueba Kuder-Richardson.

TABLA 15. VALORES DE LA PRUEBA KUDER-RICHARDSON 20 EN UNIDAD 2

k	36
$\square p_{iqi}$	0.45
\square	32.1
sx2	11.65
rkk	0.989

Se observa que la prueba arrojó un coeficiente de confiabilidad de .989, que representa un grado importante de acuerdo entre jueces, e indica que no existen diferencias significativas entre sus opiniones respecto a la pertinencia de los reactivos.

3.4.1. Análisis de reactivos del instrumento Unidad 3

El instrumento de la Unidad 3 contaba con 38 reactivos, mismos que fueron sometidos al análisis de calibración de acuerdo con la TRP (ITEMAN) y la TRI (RASCAL). Los resultados se muestran en la tabla 16.

En el análisis de acuerdo con el modelo de la TRP, realizado con el programa ITEMAN, los reactivos 14 y 36 tuvieron una correlación negativa con el resto de los reactivos (-0.19 y -0.06, respectivamente); los reactivos 14, 21 y 36 tuvieron una discriminación negativa (-0.29, -0.03 y -0.12, respectivamente). El reactivo 18 tuvo una dificultad de 1, ya que todos los participantes de la muestra lo respondieron correctamente.

TABLA 16. ANÁLISIS DE REACTIVOS DEL INSTRUMENTO DE LA UNIDAD 3

Item	1	2	3	4	5	6	7	8	9	10	11	12	13
Dif.	0.85	0.80	0.24	0.48	0.98	0.79	0.76	0.58	0.89	0.79	0.70	0.56	0.50
Disc.	0.15	0.40	0.00	0.59	0.04	0.29	0.39	0.65	0.24	0.19	0.22	0.22	0.39
Corr.	0.18	0.43	0.10	0.42	0.23	0.25	0.31	0.57	0.41	0.27	0.25	0.13	0.30
Dific. TRI	-1.19	-0.86	1.90	0.72	-3.69	-0.76	-0.58	0.32	-1.62	-0.76	-0.25	0.39	0.65
Error Est.	0.35	0.32	0.30	0.26	1.01	0.31	0.30	0.26	0.41	0.31	0.28	0.26	0.26

Item	14	15	16	17	18	19	20	21	22	23	24	25
Dif.	0.21	0.62	0.62	0.77	1.00	0.41	0.59	0.62	0.39	0.55	0.65	0.89
Disc.	-0.29	0.49	0.44	0.08	0.00	0.36	0.28	-0.03	0.46	0.23	0.34	0.12
Corr.	-0.19	0.32	0.40	0.12		0.31	0.24	0.10	0.40	0.19	0.33	0.18
Dific. TRI	2.09	0.11	0.11	-0.67	borr	1.06	0.25	0.11	1.13	0.45	-0.03	-1.62
Error Est.	0.31	0.26	0.26	0.30		0.26	0.26	0.26	0.26	0.26	0.27	0.41

Item	26	27	28	29	30	31	32	33	34	35	36
Dif.	0.44	0.52	0.42	0.35	0.50	0.77	0.47	0.91	0.62	0.67	0.67
Disc.	0.52	0.42	0.59	0.25	0.31	0.21	0.14	0.11	0.31	0.44	-0.12
Corr.	0.38	0.38	0.57	0.17	0.25	0.16	0.17	0.20	0.23	0.43	-0.06
Dific. TRI	0.92	0.59	0.99	1.34	0.65	-0.67	0.79	-1.79	0.11	-0.10	-0.10
Error Est.	0.26	0.26	0.26	0.27	0.26	0.30	0.26	0.43	0.26	0.27	0.27

Por otro lado, al aplicar el análisis de acuerdo con el modelo de TRI, realizado con el programa RASCAL, los índices de discriminación fueron adecuados, excepto en el reactivo 5, con un valor extremo de -3.69, que lo hacía demasiado fácil.

3.4.2. Ajustes al instrumento de la Unidad 3 después del análisis

Dados los datos indicados anteriormente, se decidió eliminar los siguientes reactivos: 3, 5, 14, 18, 21 y 36. Por lo tanto, este instrumento se conformó de 31 reactivos.

3.5. Resumen de resultados de todos los instrumentos analizados

Como se indica en los análisis presentados antes, los datos de jueceo fueron favorables. Un resumen de los valores obtenidos en los análisis de confiabilidad de los juecos se encuentra en la tabla 17 se muestran los resultados de la prueba Kuder-Richardson para todos los instrumentos evaluados. Estos datos fueron satisfactorios.

TABLA 17. TABLA RESUMEN DE CONFIABILIDAD EN JUECO

Instrumento	Pre-post	Unidad 1	Unidad 2	Unidad 3
kkrr	.83	.97	.97	.99

TABLA 18. REACTIVOS ACEPTADOS DESPUÉS DEL ANÁLISIS DE CALIBRACIÓN

Instrumento	Pre-post	Unidad 1	Unidad 2	Unidad 3
Núm. reactivos	25	40	34	31
Eliminados	2	8	6	5
Reactivos que integran el instrumento:	1, 2, 3, 4, 5, 6, 7, 8, 10, 11, 12, 13 15, 16, 17, 18, 19 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30	1, 2, 3, 4, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 26, 28, 29, 30, 31, 33, 34, 35, 38, 39, 41, 42, 43, 44, 45, 47, 48	1, 2, 4, 6, 7, 8, 9, 10, 11, 12, 13, 15, 16, 17, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 37, 38, 40	1, 2, 3, 4, 6, 7, 8, 9, 10, 11, 12, 13, 15, 16, 17, 19, 20, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35

Por otro lado, los resultados de las calibraciones se incluyen en la tabla 18, en la que se muestran los reactivos seleccionados en cada instrumento.

Con estos análisis se tuvieron instrumentos que cumplieron con criterios metodológicos que permiten inferir su validez y confiabilidad, con el objeto de utilizarlos en el proceso instruccional del curso descrito.

4. DISCUSIÓN

La metodología de construcción de instrumentos a utilizar en el sistema instruccional en línea se basó en el análisis cognitivo de tareas, e incorporó como dimensiones de análisis: la complejidad cognitiva de los procesos a evaluar, los modelos mentales a construir por parte de los estudiantes, y las objetivos de las unidades de aprendizaje de los programas del curso que se seleccionó para este trabajo.

Los datos de jueceo de los instrumentos confirman su validez y la pertinencia de los reactivos en términos de su apego a los contenidos y objetivos del programa. Por otro lado, los análisis de calibración permitieron, desde las teorías psicométricas clásica y de respuesta al ítem, seleccionar los reactivos que permitían identificar de una forma más confiable los conocimientos de los estudiantes y descartar reactivos dada su falta de discriminación o correlación, o bien por tener un nivel de dificultad extremo.

Las implicaciones de estos resultados se relacionan con la importancia de contar con una metodología para la construcción de instrumentos de evaluación objetiva que permita realizar diagnósticos de los niveles de desempeño de los estudiantes en contextos académicos como el que se describe en el presente trabajo. Especialmente en el contexto de la educación mediante Internet, en la que es preciso construir ambientes que confieran a la evaluación la función doble: de apoyo al aprendizaje, y al mismo tiempo de diagnóstico del desempeño. El presente trabajo contribuye en la medida en que presenta esquemas de evaluación útiles para valorar en la investigación los impactos de una diversidad de componentes incluidos en los ambientes de aprendizaje en línea, lo cual es importante ante la escasez de estudios que realmente permiten valorar los elementos que determinan el desempeño en estos contextos (Sadik, 2001), y ante la consideración de que la evaluación es un "punto débil" de la educación virtual (Barberá, 2006).

El modelo que se presenta adapta una metodología probada en el área de la evaluación del aprendizaje complejo en la educación superior presencial, para incorporarla a los ambientes virtuales como un mecanismo que puede dar estructura a las evaluaciones en un ambiente de aprendizaje en línea, partiendo de la "disección" de las competencias. De esta forma, las competencias y sus componentes, que representan las metas de aprendizaje de los programas educativos, son el eje de la evaluación.

Una vez analizados los contenidos e identificadas las competencias del dominio educativo, el Análisis Cognitivo de Tareas conduce a un potente esquema de evaluación que permite diagnosticar la estructuración del conocimiento de los estudiantes a lo largo de cursos en línea, tomando en cuenta dimensiones centrales de sus procesos cognitivos. La creación del modelo del dominio de conocimiento de los cursos conduce a la construcción de instrumentos de evaluación válidos y confiables.

Modelar estas estructuras puede ser útil en la investigación del aprendizaje en línea, ya que ayuda a determinar la influencia de variables. Investigaciones de este tipo se requieren en el campo de la educación a distancia (Jegede, 1999), y en línea (Quesada, 2006; Barberá, 2006). La ausencia de

investigación acerca de la evaluación del aprendizaje es evidente, y propuestas de este tipo podrían ayudar a generar conocimiento y métodos de estudio en el área.

En suma, la presente propuesta podría aplicarse para realizar investigación acerca de diversos métodos de evaluación y fomento del aprendizaje en línea, gracias a la seguridad que ofrece en el sentido de delimitar el dominio educativo bajo estudio mediante una taxonomía válida y sensible.

La educación en línea es considerada una opción flexible, efectiva y en tanto viable para cubrir la alta demanda educativa de nuestro país, pero si no generamos investigación que explique los efectos de nuestras intervenciones de evaluación y fomento del aprendizaje en estos contextos, en este pragmatismo podríamos dirigirnos hacia un escenario caracterizado por malos resultados, que es lo que ya no queremos en la educación mexicana.

REFERENCIAS BIBLIOGRÁFICAS

- Astolfi, J.P. (2003). *Aprender en la escuela*. Santiago de Chile: LOM.
- Anderson, T. (2004). Toward a theory of online learning. En T. Anderson y F. Elloumi (Eds), *Theory and practice of online learning* (pp. 33-58). Athabasca: Athabasca University.
- Barberá, E. (2006). Aplicaciones de la tecnología a la e-Evaluación. *Revista de Educación a Distancia*, 5, Consultado el 27 de julio de 2008 en: <http://www.um.es.ead/red/M6>.
- Bransford, J., Brown, A. L., y Cockings, R. R. (2004). *How people learn: Brain, mind, experience, and school. Expanded edition*. Washington, D.C.: National Academy Press.
- Brown, F. (1980). *Principios de la medición en Psicología y Educación*. Bogotá: Manual Moderno.
- Castañeda, S. (1998a). *Evaluación y fomento del desarrollo intelectual en la enseñanza de ciencias, artes y técnicas: Perspectiva internacional en el umbral del siglo XXI*. México: Editorial Miguel Angel Porrúa - UNAM.
- Castañeda, S. (1998b). Evaluación de resultados de aprendizaje en escenarios educativos. *Revista Sonorense de Psicología*, 12, pp. 57-67.
- Castañeda, S. (2002). A cognitive model for learning outcomes assessment. *International Journal of Continuing Engineering Education and Life-long Learning*, 12, pp. 94-106.
- Castañeda, S. (2004a). Educación, aprendizaje y cognición. En S. Castañeda, S. (Ed). *Educación, aprendizaje y cognición: teoría en la práctica* (pp. 49-74). México: Manual Moderno.
- Castañeda, S. (2004b). Enseñanza estratégica: guía abreviada para el docente y el tutor En S. Castañeda, S. (Ed), *Educación, aprendizaje y cognición: teoría en la práctica* (pp. 349-421). México: Manual Moderno.
- Castañeda, S. (2006). Evaluación del aprendizaje en educación superior. En: S Castañeda (coord). *Evaluación del aprendizaje en el nivel universitario: elaboración de exámenes y reactivos objetivos* (pp. 3-27). México: UNAM.

- Chipman, S. F., Schraagen, J. M., Shalin, V. L. (2000). Introduction to Cognitive Task Analysis. En J. M. Schraagen, S. F. Chipman y V. L. Shalin (eds.), *Cognitive Task Analysis* (pp. 3-23). New York: Lawrence Erlbaum Associates.
- Hambleton, R.K., Rogers, H.J. y Swaminathan, H. (1991). *Fundamentals of item response theory*. Londres: Sage Publications.
- ITEMAN (1993). *Conventional Item Analysis Program*. (Versión 3.5). St. Paul, MN: Assessment Systems Corp.
- López, M. (2006a). Preparación y calibración de reactivos: medición de y para el aprendizaje. En S. Castañeda (Coord.) *Evaluación del Aprendizaje en el Nivel Universitario: Elaboración de exámenes y reactivos objetivos* (pp. 29-56). México: UNAM.
- López, O. M. (2006b). Análisis de reactivos. En S. Castañeda (Coord.), *Evaluación del Aprendizaje en el Nivel Universitario: Elaboración de exámenes y reactivos objetivos* (pp. 215-245) México: UNAM.
- Peñalosa, E. y Castañeda, S. (2008a). Generación de conocimiento en la educación en línea: un modelo para el fomento de aprendizaje activo y autorregulado. *Revista Mexicana de Investigación Educativa*, 13 (36), 229-260.
- Peñalosa, E. y Castañeda, S. (2008b). Meta-Tutor: an online environment for knowledge construction and self-regulated learning in clinical psychology teaching. *International Journal of Continuing Education, Engineering and Lifelong Learning*, 18 (3), pp. 283-297.
- Quesada, R. (2006). Evaluación del aprendizaje en la educación a distancia en línea. *Revista de Educación a Distancia*, 5, Consultado el 27 de julio de 2008 en <http://www.um.es/ead/red/M6>.
- RASCAL (1992). *Rasch Analysis Program* (Versión 3.5). St. Paul, MN: Assessment Systems Corp.
- Renkl, A., Mandl, H. y Gruber, H. (1996). Inert knowledge: Analyses and remedies. *Educational Psychologist*, 3, pp. 115-121.
- Sadik, A. (2001). *Directions for future research in on line distance education*. Documento de disertación doctoral, Facultad de Educación de la Universidad South Valley, Qena 11183, Egipto.
- Schwartz, D.L., Brophy, S., Lin, X. y Bransford, J.D. (1999). Software for managing complex learning: examples from an educational psychology course. *Educational Technology Research and Development*, 47, pp. 39-60.
- Van Merriënboer, J.G., Clark, R. y deCrook, M.B.M (2002). Blueprints for complex learning: the 4 C/ID – Model. *Education, Training and Development*, 50, pp. 39-64.



ISSN: 1989-0397

LA EVALUACIÓN DEL LOGRO ESCOLAR EN LA EDUCACIÓN BÁSICA EN MÉXICO. LOS EXCALE DE MATEMÁTICAS

Andrés Sánchez

Revista Iberoamericana de Evaluación Educativa 2009 - Volumen 2, Número 1

<http://www.rinace.net/riee/numeros/vol2-num1/art10.pdf>

Fecha de recepción: 15 de marzo de 2009

Fecha de comunicación de dictamen: 18 de abril de 2009

Fecha de aceptación: 26 de abril de 2009

El Instituto Nacional para la Evaluación de la Educación (INEE) fue creado en México por decreto presidencial el 8 de agosto de 2002. Su existencia obedece a la necesidad de contar en el país con evaluaciones externas (esto es, realizadas por instancias independientes de la Secretaría de Educación Pública) del Sistema Educativo Nacional en su conjunto.

Entre otras actividades, el INEE es el encargado de coordinar las pruebas de logro educativo internacionales de la OCDE (PISA) y de la Oficina Regional de Educación de la UNESCO para América Latina y el Caribe en México, edita anualmente el Panorama Educativo de México en el que se incluye un conjunto de indicadores educativos organizados coherentemente, lleva a cabo estudios sobre recursos y procesos educativos y diseña, construye, aplica y analiza las pruebas nacionales Excale (INEE, 2005a). Todo esto para ofrecer a los tomadores de decisiones información de buena calidad que permita conocer y comprender de mejor manera la educación básica y media superior¹ del país.

Unas de las líneas fundamentales de trabajo del INEE, la de Evaluación de la calidad del sistema educativo, incluye la sub línea de Evaluación de los resultados educativos, centrándose en los resultados del aprendizaje de los alumnos, aunque reconociendo que los productos de la educación son más amplios (INEE, 2006).

Para evaluar los resultados del aprendizaje de los alumnos con referencia a lo que el currículo nacional indica que debe aprenderse, se han desarrollado desde 2004 las pruebas Excale (Exámenes de la Calidad y el Logro Educativos), aplicadas por primera vez en junio de 2005. Los resultados de los Excale ayudan a construir una visión general de lo que los estudiantes aprenden como resultado de su escolarización formal; conocer los puntos débiles y fuertes del currículo nacional, atendiendo variables como grado, asignatura, modalidad educativa, género, entidad y establecer comparaciones en el tiempo que pueden mostrar tendencias de aprendizaje nacionales y desagregadas a partir de variables como las señaladas. No se ofrecen resultados por alumno, ni por grupo, ni por escuela, dado que los Excale están diseñados para la evaluación del Sistema Educativo Nacional en su conjunto.

Estas pruebas se aplican a niños de tercero de preescolar (que tienen alrededor de cinco años de edad), tercero de primaria (alrededor de 8 años), sexto de primaria (alrededor de 11 años) y tercero de secundaria (alrededor de 14 años), en las asignaturas de Español, Matemáticas, Ciencias naturales y Ciencias sociales². La elección de las áreas curriculares que se evalúan con las pruebas Excale obedece a la necesidad de limitar el marco de trabajo. Todas las asignaturas que establece el currículo nacional tienen importancia para el desarrollo de los estudiantes y por lo tanto son susceptibles de ser tomadas en cuenta para desarrollar una visión de conjunto que permita conocer el estado de la educación en México; sin embargo, ante la imposibilidad de desarrollar simultáneamente pruebas de buena calidad de todas las asignaturas, se decidió elegir las instrumentales (Español y Matemáticas), que por ayudar al desarrollo de lenguajes permiten al estudiante seguir aprendiendo a partir de libros y otros medios de información, y las que agrupan grandes áreas curriculares, como son las Ciencias naturales y las Ciencias sociales.

Actualmente, ya se han realizado cuatro aplicaciones de las pruebas Excale. Los resultados de cada aplicación se presentan en una publicación anual llamada *La Calidad de la Educación Básica en México*.

¹ Estos dos niveles son de la competencia del INEE. No así la educación superior.

² Las Ciencias naturales en secundaria, y las Ciencias sociales tanto en primaria como en secundaria, más que ser asignaturas son agrupaciones de asignaturas con elementos comunes. No todas las áreas temáticas de estas dos agrupaciones han sido evaluadas en cada grado escolar que se señala.

Además, se producen reportes puntuales para analizar algunas de las características de los resultados, como *Factores escolares y aprendizaje en México. El caso de la educación básica* (Backhoff y col. 2007c). También se hacen públicas las bases de datos completas de cada aplicación, junto con la documentación necesaria para utilizarlas para hacer análisis específicos. Todo esto permite obtener elementos para la reflexión en torno a los resultados, ya sea de manera general o a partir de selecciones de los resultados para analizar diferentes temáticas, como podrían ser los distintos estratos educativos muestreados, los diferentes años de aplicación, el género de los participantes, una entidad en particular o una asignatura. Este último caso es el que corresponde con este escrito, cuyo objetivo es hacer una breve nota a “vista de pájaro” sobre el logro educativo en Matemáticas de años recientes en educación básica, en México, a partir de los resultados de las pruebas Excale. En primer lugar se mostrará el plan de pruebas que existe para la aplicación de los Excale a lo largo de los años. En un segundo apartado se describen las características principales de los Excale de Matemáticas. El tercer apartado explica brevemente cómo se construyen y aplican las pruebas Excale. En el cuarto, se muestran algunos resultados de cada una de las aplicaciones que se han hecho desde 2005 de las pruebas Excale de Matemáticas³. En el quinto apartado se ofrecen las conclusiones de trabajo.

1. DISTRIBUCIÓN DE LA APLICACIÓN DE PRUEBAS EXCALE EN EL TIEMPO

Una característica interesante de las pruebas Excale es que no se aplican a todos los grados y asignaturas que son de su competencia cada año. Las pruebas se dosifican en un arreglo basado en los ciclos escolares, de tal modo que una serie de pruebas se aplican cuando el año escolar está a punto de concluir, sólo a una muestra representativa de la población y sólo refiriéndose a algún grado escolar. El siguiente esquema muestra el plan de aplicación de pruebas que se tiene hasta el año 2016.

TABLA 1. PLAN DE APLICACIÓN DE PRUEBAS EXCALE PARA EDUCACIÓN BÁSICA

Grados	Años escolares											
	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016
3° Preescolar			E, M				E, M				E, M	
3° Primaria		E, M N, S				E, M N, S				E, M N, S		
6° Primaria	E, M		E, M		E, M N, S		E, M		E, M N, S		E, M	
3° Secundaria	E, M			E, M N, S		E, M		E, M N, S		E, M		E, M N, S

Tomada de Backhoff y col. 2005:20.

En la tabla, E significa Español; M matemáticas; N Ciencias naturales y S Ciencias sociales. Puede observarse, por ejemplo, que en 2007 se aplicó la prueba de tercero de preescolar de Español (Lenguaje y comunicación) y Matemáticas (Pensamiento matemático). El arreglo no es caprichoso; un mismo grado se

³ En realidad, hay una aplicación de Excale que no se está incluyendo en el estudio. Es la réplica del Excale de sexto de primaria que se realizó dos años después, en 2007, con una versión idéntica de la prueba. En términos de los resultados que se están mostrando aquí, se considera que incluir dicho estudio no agregaría información y sí aumentaría la complejidad de lo reportado, por lo que se opta por dejarlo fuera. De cualquier modo, en la bibliografía se incluye la referencia al reporte extenso del estudio comparativo 2005-2007 al que dio origen la réplica.

evalúa cada cuatro años, para poder hacer comparaciones horizontales en el tiempo; se puede hacer un seguimiento de cohortes a partir de 2007, pues las aplicaciones coinciden con una misma generación de estudiantes que va cambiando de grado escolar; finalmente, puede tomarse un punto cualquiera tras una serie de aplicaciones y mirar lo que se sabe de los diferentes grados y asignaturas, que es precisamente lo que se hace en la parte de resultados de este trabajo: considerar los resultados de los Excales reportados hasta el 2008, incluyendo los resultados de sexto de primaria y tercero de secundaria (2005), tercero de primaria (2006) y tercero de preescolar (2007).

2. LAS PRUEBAS EXCALE DE MATEMÁTICAS

Los Excale de Matemáticas, al igual que las de otras asignaturas, tienen tres características fundamentales:

- Son criterios, es decir, utilizan un referente que no se relaciona directamente con la ejecución de los demás individuos que realizan la prueba, sino con criterios establecidos previamente.
- Están alineados al currículo nacional; precisamente es lo que el currículo indica que debería saberse al finalizar determinado grado educativo lo que sirve como criterio de la prueba.
- Son matriciales en cuanto a su diseño, lo cual significa que, ante una cantidad de preguntas muy grande para describir suficientemente la escala, se presenta a cada niño solamente una fracción de ésta, para evitar el efecto de la fatiga en las respuestas. La parte de las preguntas que contesta cada niño se traslapa con la parte que contestan los demás niños, de modo que las respuestas de un grupo grande de niños, debidamente analizadas, permite saber qué contenidos dominan como grupo; no así lo que sabe cada niño (pues contestaron diferentes partes de la prueba extensa). Esto es útil para pruebas como el Excale, que no tienen la finalidad de dar resultados por alumno, o por escuela, sino solamente por grandes agregaciones de estudiantes: a nivel estatal, nacional, o por estrato educativo.

Además, los Excale de Matemáticas tienen algunas particularidades con respecto a otras pruebas: todas sus preguntas son de opción múltiple con cuatro opciones, no se requiere calculadora u otras herramientas para resolverse, además de un lápiz.

Por otra parte, la estructura de los Excale de matemáticas está basada en líneas de evaluación. Las líneas de evaluación son contenidos relacionados en términos de relaciones de servicio, es decir, un contenido más simple sirve como elemento facilitador para el aprendizaje de otro más complejo, facilitando su comprensión o ejecución. Si un estudiante es capaz de resolver un reactivo que implica conocimientos y habilidades de varios tipos (por ejemplo, el algoritmo de la suma, el de la resta, la interpretación correcta de la solicitud del problema y sus datos), sabemos que puede operar con todos estos conocimientos y habilidades; pero cuando un estudiante responde incorrectamente a un reactivo cuya solución implica todos estos conocimientos y habilidades, no tenemos claro qué es lo que no sabe. Pudiera ser que no domine ninguno de los conocimientos y habilidades que se ponen en juego en el reactivo, o que no sepa sólo uno de ellos, así como cualquier otra posibilidad intermedia. Las pruebas Excale, además de reactivos que implican varios conocimientos y habilidades, incluyen reactivos más simples, que combinan menos conocimientos y habilidades con respecto al reactivo complejo (por ejemplo, un reactivo que incluyera solamente el algoritmo de la suma, sin contexto ni datos que interpretar), de manera que pueda

analizarse hasta qué punto se han adquirido los contenidos curriculares, cuyos elementos finales tienden a ser integraciones de los contenidos previos, para el caso del currículo mexicano de Matemáticas en Educación básica.

Veamos un ejemplo ilustrativo, sacado de la prueba de sexto de primaria. Cuando la tarea requiere resolver problemas que impliquen una suma de fracciones, un 26 % de los estudiantes logra realizarla. ¿Qué pasa con las tres cuartas partes restantes de los estudiantes? ¿qué es lo que no saben hacer y, tal vez más importante, ¿qué es lo que sí saben hacer? Bueno, si la tarea requiere sumar dos fracciones con diferente denominador pero no en el contexto de un problema, 30% de los estudiantes logran resolverla. Si se requiere sumar fracciones, pero con el mismo denominador sin el contexto de un problema, el 52 % de los estudiantes lo realizan correctamente. Resolver problemas que impliquen una suma de números naturales en su significado de unir es algo que el 67% de los alumnos logran (datos obtenidos de la tabla que se encuentra en Backhoff y col., 2006:295-297).

Así, a partir de una línea de evaluación que va desde la resolución de problemas más sencillos (en este caso, que impliquen sumar números enteros), hasta la resolución de problemas que incorporan más elementos que deben conocerse (aquí, que impliquen sumar números fraccionarios), podemos darnos cuenta de los elementos que causan mayores dificultades (en este caso, el uso de fracciones con distinto denominador), y los elementos que son dominados por la mayor parte de los estudiantes (en el ejemplo, el reconocer que el problema implica sumar los elementos).

3. CONSTRUCCIÓN Y APLICACIÓN DE PRUEBAS

Los Excale tienen una metodología rigurosa para su proceso de diseño, construcción, aplicación y validación, adaptando el modelo planteado por Nitko (1994) para desarrollar pruebas nacionales. Un elemento fundamental de dicha metodología es el trabajo colegiado de especialistas en educación incluyendo profesores frente a grupo, directores, investigadores educativos, autores de libros de texto y diseñadores de currículo y autoridades educativas. El trabajo colegiado se organiza en comités sucesivos que tienen tareas específicas, y que basan su trabajo en los productos de los comités anteriores. Para ser breves, se presenta en la tabla 2 un resumen del proceso (éste se encuentra desglosado en Backhoff y col. 2005).

TABLA 2. RESUMEN DEL PROCESO DE DISEÑO, CONSTRUCCIÓN Y VALIDACIÓN DE LOS EXCALE

Comité	Proceso	Producto
Académico	Responder a la pregunta ¿qué del currículo vigente debe ser evaluado en las pruebas Excale?	Retícula que relaciona los contenidos curriculares. Tabla de contenidos.
De Especificaciones	Responder a la pregunta ¿cómo deben hacerse la evaluación de los contenidos curriculares propuestos para ello por el comité anterior?	Especificaciones de reactivos, que son descripciones detalladas de cómo deben ser los reactivos para evaluar un contenido dado.
De Reactivos	Redactar una serie de reactivos específicos para cada contenido, siguiendo las guías de las especificaciones.	Reactivos asociados a las especificaciones y a la Tabla de contenidos.
De Validez y sesgo	Verificar que los reactivos del comité anterior sean pertinentes para evaluar el currículo, su redacción adecuada y estén libres de sesgo cultural y de género.	Observaciones a los reactivos que no cumplen con los criterios de validez.

Tras estos procesos, los reactivos son piloteados y a partir de los resultados del piloteo se hacen los ajustes necesarios a la prueba para su aplicación. La aplicación definitiva se hace a muestras

representativas de la población nacional, y en muchas ocasiones a muestras representativas a nivel de las 32 entidades en que se divide el país. Las muestras también representan a los diferentes estratos educativos de los que se hablará en la sección de Resultados y análisis. Los resultados que se ofrecen en la siguiente sección no son los resultados de la muestra, sino los estimados para la población, que tienen un error de aproximadamente $\pm 1\%$.

Las puntuaciones de las pruebas Excale, incluyendo las de Matemáticas, se calculan utilizando el modelo de Rasch y estableciendo una escala que va de 200 a 800 puntos, con una media de 500 y una desviación estándar de 100. Las aplicaciones sucesivas de una misma escala utilizan los parámetros originales, por lo que su media se desplaza del 500 cuando hay variaciones en el aprendizaje, o cuando existen errores de medición importantes. Para el segmento de los resultados utilizado en este trabajo, no se utiliza la escala descrita, debido a que se prefirió analizar los resultados en términos de la cantidad de estudiantes que se encuentran en los niveles de logro definidos para estas pruebas. Desde nuestro punto de vista, conocer y comparar niveles de logro aporta mucho más a la reflexión que conocer y comparar puntos únicos en una escala, por las siguientes razones:

1. Cada nivel de logro es definido en términos de lo que los niños de determinada población que se encuentran en él pueden hacer, de tal manera que al hablar del porcentaje de estudiantes que se encuentra en cada nivel de logro se está hablando de un porcentaje de estudiantes que sabe hacer tales y cuales cosas, pero no tales y cuales otras, permitiendo comprender con mayor claridad el estado de la educación en esa población. La puntuación promedio no ofrece la riqueza de esta información.
2. Además, esta distribución permite que los especialistas en la asignatura reportada obtengan, en su análisis, un mapa de los contenidos que se dificultan más a los estudiantes, los que prácticamente toda la población domina y los que están en niveles intermedios de adquisición. Nuevamente, esta información no se ofrece al utilizar sólo el promedio de la puntuación.
3. Al establecer cuatro niveles de logro, los análisis en el tiempo pueden mostrar cambios en sólo algunos de estos niveles, o cambios asimétricos entre niveles, que nos permitan hacer hipótesis sobre movimientos de estudiantes de un nivel a otro (por ejemplo, si en una segunda aplicación tras algunos años se encuentra que el porcentaje de alumnos en el nivel más bajo disminuyó en un 15 % y aumentó en la misma proporción en el segundo nivel, es un caso diferente que si el porcentaje de alumnos en el nivel más bajo disminuye en un 15 % pero el porcentaje de alumnos en el segundo nivel aumenta sólo un 7 %) . El uso simple del puntaje promedio no permite ver estas dinámicas.
4. Finalmente, utilizar niveles de logro nos permite abonar en este trabajo hacia la necesidad de reflexión sobre los contenidos que se adquieren por encima de la reflexión sobre los puntajes que se obtienen.

Las pruebas Excale han ofrecido sus resultados desde 2006 de varias maneras, buscando no sobresimplificar la realidad educativa nacional (lo cual podría ocurrir si se reportaran sólo promedios y otros datos numéricos desligados de información cualitativa). Por las razones mencionadas arriba, uno de los formatos, que nos parece particularmente útil para examinar los resultados educativos en conjunto, es el uso de niveles de logro. En el INEE utilizamos cuatro niveles de logro, que tienen una expresión general, que puede leerse en la tabla 3, y una expresión particular a cada grado y asignatura.

TABLA 3. NIVELES GENERALES DE LOGRO PARA EXPRESAR LOS RESULTADOS DE LAS PRUEBAS EXCALE

Niveles de logro	Competencias académicas
Avanzado	Indica un dominio muy avanzado (intenso, inmejorable, óptimo o superior) de conocimientos, habilidades y destrezas escolares que reflejan el aprovechamiento máximo de lo previsto en el currículo.
Medio	Indica un dominio sustancial (adecuado, apropiado, correcto o considerable) de conocimientos, habilidades y destrezas escolares que pone de manifiesto un buen aprovechamiento de lo previsto en el currículo.
Básico	Indica un dominio imprescindible (suficiente, mínimo, esencial, fundamental o elemental) de conocimientos, habilidades y destrezas escolares necesarios para poder seguir progresando satisfactoriamente en la materia.
Por debajo del básico	Indica carencias importantes en el dominio curricular de los conocimientos, habilidades y destrezas escolares que expresan una limitación para poder seguir progresando satisfactoriamente en la materia.

Tomada de Backhoff y col. (2007).

Los niveles de logro son establecidos utilizando el procedimiento conocido como Bookmark, en el que un comité de especialistas "define las categorías y niveles de logro de los estudiantes de acuerdo a 1) la ejecución que presuntamente se esperaría de ellos y 2) los resultados reales en la prueba respectiva" (Backhoff y col., 2005:18).

4. RESULTADOS Y ANÁLISIS

Tomando en cuenta lo anterior, se presentarán resultados de las pruebas Excale de Matemáticas de tercero de secundaria (2005), sexto de primaria (2005), tercero de primaria (2006) y tercero de preescolar (2007), en términos de los niveles de logro que alcanzan los niños que se encuentran en estos grados escolares. La intención fundamental es la de mostrar una breve panorámica de lo que la mayoría de los niños de estos grados es capaz de hacer con respecto a lo que el currículo de matemáticas indica que deberían hacer. Ya se ha visto, siguiendo el plan de pruebas, que los Excale no han sido aplicados en orden cronológico ascendente por grados escolares. Los dos primeros Excale que se aplicaron fueron los de sexto de primaria y tercero de secundaria.

Considerando cuál sería la mejor manera de presentar los resultados de varios años de aplicación de los Excale de matemáticas para este trabajo, una primera idea fue la de presentar los resultados en orden por su año de aplicación. No obstante, parece mejor mostrar los resultados siguiendo el orden ascendente de los grados escolares, de manera que pueda percibirse el aumento de exigencia de las escalas, así como los cambios en la distribución de los resultados. Debido a este cambio, es importante aclarar que los resultados de los diferentes grados escolares no se refieren a una misma población, ni a muestras consecutivas tomadas de una misma generación de estudiantes. Otra decisión que se ha tenido que tomar para la presentación de resultados en este trabajo es la de elegir algunos de los resultados reportados extensamente en los informes institucionales. Con la intención de utilizar algunos de los más elocuentes, nos hemos decidido por utilizar fundamentalmente el porcentaje de estudiantes que se encuentra en los cuatro niveles de logro previamente mencionados.

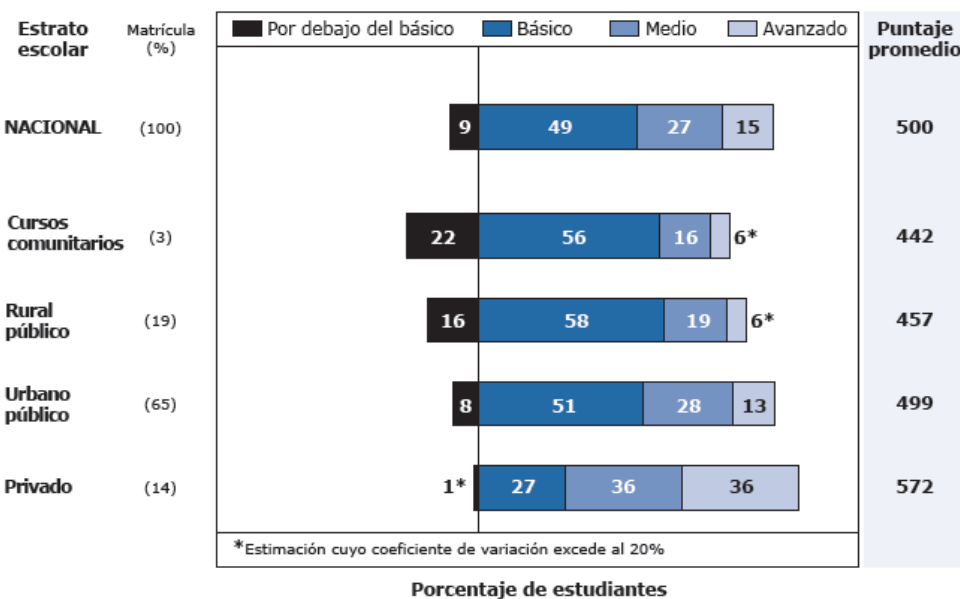
Debido a que sería engorroso en este breve espacio enumerar todos los niveles de logro particulares de los diferentes grados evaluados con los Excale de Matemáticas, hemos decidido incluir en este trabajo sólo la descripción del nivel en que se encuentre la mayoría de alumnos de un grado específico, a manera de ilustración. Las descripciones completas, con ejemplos, de todos los niveles de logro específicos pueden encontrarse en Backhoff y col. (2006), (2007) y (2008). En estos mismos documentos pueden encontrarse datos sobre la evaluación de otras asignaturas, así como información más pormenorizada de las pruebas utilizadas y de los resultados obtenidos.

4.1. Tercero de preescolar

La evaluación del aprendizaje en este grado está alineada al Programa de Educación Preescolar 2004, que es el currículo de aplicación nacional actualmente. Para el área de Pensamiento matemático, este programa incluye ocho competencias, divididas en dos aspectos: Número y Forma, espacio y medida. El Excale de preescolar incluye tareas que evalúan aspectos relevantes de cada una de ellas.

En la ilustración 1 puede observarse el porcentaje de alumnos en el país que se encuentra en cada uno de los cuatro niveles de logro. También se muestran los porcentajes en que se encuentran los niños de cuatro estratos educativos que atienden a porcentajes elevados de la matrícula: Cursos comunitarios, Escuela rural pública, Escuela urbana pública y Escuela Privada⁴.

ILUSTRACIÓN 1. PORCENTAJE DE ALUMNOS POR NIVEL DE LOGRO EDUCATIVO Y ESTRATO ESCOLAR EN EXCALE DE PENSAMIENTO MATEMÁTICO (TERCERO DE PREESCOLAR, APLICACIÓN 2007)



Tomada de Backhoff y col. (2008).

⁴ Los estratos educativos se definen por el tipo de escuela en que se toma determinada parte de la muestra: los cursos comunitarios son escuelas de sostenimiento público ubicadas en localidades de difícil acceso y escasa población, donde no existen servicios educativos regulares. Las escuelas rurales públicas son escuelas de sostenimiento público ubicadas en comunidades con una población menor a 2,500 habitantes. Las escuelas urbanas públicas son como las anteriores pero en comunidades con población mayor a 2,500 habitantes. Finalmente, las escuelas privadas son las que tienen un sostenimiento privado.

Puede observarse que la mayor parte de los niños (49 %) se encuentran en el nivel básico. Ya que los niveles de logro son acumulativos, los niños de nivel Básico pueden hacer lo que se describe para el nivel Por debajo del básico y lo que se describe para el nivel Básico. Así, los niños que se encuentran en dicho nivel "son capaces de decir la serie numérica empezando por el uno, y hasta el que saben (máximo treinta); enumerar de manera oral objetos desde el uno y hasta el que saben (máximo veinte), siguiendo el orden de la serie numérica; identificar los usos de los números en situaciones cotidianas; y escribir en orden un tramo de la serie numérica convencional (máximo treinta) o números que se les dictan", "son capaces de utilizar números para representar cantidades menores de siete; contar colecciones de objetos, hasta treinta; comparar colecciones de objetos y establecer relaciones de igualdad y desigualdad; registrar la cantidad de elementos en tablas y gráficas, con ayudas como categorías establecidas y ejemplos; identificar en tablas o gráficas la colección en la que hay más o hay menos elementos; identificar la colección faltante en una serie de colecciones con patrón de crecimiento $n + 1$; construir la colección que sigue en una serie de colecciones con patrón de crecimiento $n + 1$ ", "reconocer un objeto que se les presenta gráficamente, desde distintos puntos espaciales; e identificar de manera perceptual propiedades medibles contrastantes de los objetos (lleno-vacío, alto-bajo, largo-corto...)" e "identificar semejanzas entre un cuerpo geométrico y un objeto del entorno; identificar posiciones de objetos respecto a otros objetos, en una representación gráfica; comparar de manera perceptual la longitud de objetos: más corto que..., mas largo que...; resolver problemas que impliquen estimar longitudes; y distinguir el instrumento apropiado para medir el peso" (Backhoff y col. 2007:47-48).

Por otro lado, podemos conocer también lo que los niños del nivel básico no logran adquirir del currículo: los elementos que están consignados en la descripción de los niveles de logro más elevados. Conocerlos es importante, puesto que son cosas que la mayoría de los niños del país (todos los que se encuentran en el nivel básico y también en el nivel por debajo del básico) no logran hacer. Así pues, el 58% (49% +9%) de los niños de tercero de preescolar no logran "utilizar números para representar cantidades menores de 13; identificar y representar numerales hasta treinta; emplear los números para identificar hasta la cuarta posición de un elemento en una serie ordenada; resolver problemas que impliquen combinar el valor de monedas de cincuenta centavos, uno y cinco pesos; identificar patrones diseñados en función de un criterio de repetición de dos o tres elementos; y resolver problemas que impliquen agregar, igualar, comparar, quitar o repartir cantidades de una o más colecciones de hasta nueve objetos con o sin el apoyo de la representación gráfica", ni "utilizar números para representar cantidades hasta veinte". (Op. Cit.:46-47) en cuanto al aspecto de número.

En relación con el aspecto de Forma, espacio y medida, el 58% de los niños de tercero de preescolar no 'reconocen figuras geométricas que comparten atributos; identifican figuras semejantes a una de muestra; identifican los cambios que ocurren en una figura geométrica al combinarla con otras; no resuelven problemas que impliquen medir longitudes tomando en cuenta una unidad de medida no convencional; no identifican el orden de puntos de referencia espacial en un trayecto representado gráficamente; ni establecen o identifican la secuencia de imágenes que representan las relaciones temporales antes-después-al final'; tampoco logran 'trazar trayectos a partir de puntos de referencia

espaciales que incluyen direccionalidad (desde, hacia, hasta) ni ubicar los días de la semana a partir de las actividades que realizan' (Op. Cit.:46-47).⁵

Mirar lo que pueden y no pueden hacer los niños que se encuentran en el nivel básico, sabiendo que son la mayoría a nivel nacional, tiene un sabor agri dulce: por una parte, no es desdeñable la cantidad de habilidades que los niños con este nivel de logro demuestran, sobre todo considerando que el programa de preescolar de 2004 era aún muy novedoso, en muchos casos podríamos decir desconocido, para las educadoras y las directoras de preescolar para el 2006, año en que se aplicó la prueba.

Por otra parte, hay un nueve por ciento de niños a nivel nacional que se encuentran en el nivel Por debajo del básico, que si bien en términos de porcentajes puede parecer poco, en la realidad nacional significa más de doscientos mil niños que no están adquiriendo las competencias fundamentales para seguir aprendiendo matemáticas, es decir, niños que ingresarán a la educación primaria con una desventaja inicial importante. Además, cuando examinamos a las poblaciones desfavorecidas (Cursos comunitarios y Escuelas rurales públicas) este porcentaje aumenta mucho. Además, la cantidad de contenidos curriculares que la mayoría de los niños de este grado no logra hacer es muy considerable.

Es importante enfatizar que, tanto en el análisis de resultados de este grado como en los posteriores, cuando hablamos de estratos educativos como Cursos comunitarios, Escuelas rurales públicas, Escuelas privadas, etcétera, no es posible inferir de los datos presentados que el tipo de escuela al que está asociado el estrato es causante del nivel de logro; estudios realizados por el INEE a partir de los cuestionarios de contexto que se aplican junto con los Excale han mostrado que los alumnos que asisten a estos tipos de escuelas se diferencian de quienes asisten a otros tipos en una serie de variables sociales, culturales y económicas, y que éstas influyen en mayor o menor medida en el logro escolar (Backhoff y col., 2007b; Backhoff y col., 2007c).

Queremos comentar sobre otros resultados de los Excale que nos parecen especialmente interesantes y útiles para comprender el aprendizaje del logro matemático en la educación básica en México, aunque en este espacio no será posible más que presentarlos de manera fragmentada, y sin mayor análisis: cada reporte extenso de resultados Backhoff y col. (2005), (2006) y (2007) incluye tablas en que se describe cada reactivo de los Excale en términos de lo que evalúa y se indica la dificultad relativa de cada reactivo con respecto a los demás de la escala, así como el porcentaje de aciertos por estrato educativo y a nivel nacional. Esta información es útil para tener elementos sobre el dominio que tienen los estudiantes de los diferentes contenidos escolares. En ocasiones, el nivel de dificultad de algunos contenidos es muy distinto al que supondría la intuición y la experiencia de los profesores, los diseñadores de currículo y otros lectores de los resultados; en estos casos la información sistematizada de los Excale es especialmente útil. La tabla 4 es un fragmento, a modo de ejemplo, de este esquema para los reactivos de Pensamiento matemático de preescolar. Existen datos similares para los otros grados evaluados por los Excale de Matemáticas.

⁵ Se usa una sola comilla en la cita, indicando que no es exacta, debido a que los tiempos verbales se cambiaron para facilitar la comprensión de que estamos hablando de lo que los niños no saben hacer, ya que en el original citado se habla de lo que sí pueden hacer los alumnos de los niveles de logro superiores al básico.

**TABLA 4. PORCENTAJE DE ACIERTOS POR REACTIVO (INDICADOR).
EXCALE DE PENSAMIENTO MATEMÁTICO, PREESCOLAR (FRAGMENTO)**

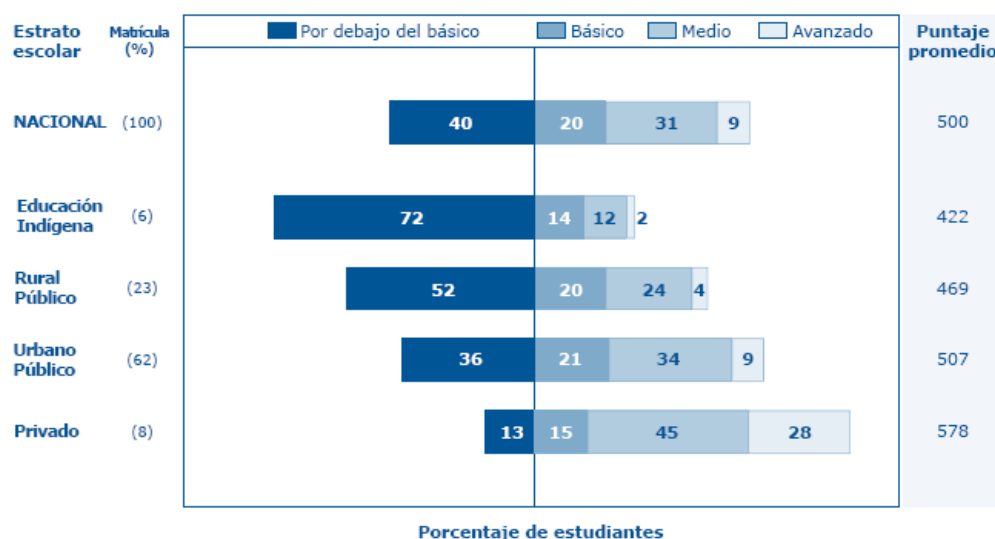
Reactivo	Indicador	Dificultad	Porcentaje de aciertos				
			Nacional	Cursos comunitarios	Rural público	Urbano público	Privado
Utiliza los números en situaciones variadas que implican poner en juego los principios del conteo							
1 ^A	Dice la serie numérica empezando por el uno, y hasta el que sabe (máximo treinta)	181	97	97	98	97	98
2 ^A	Enumera de manera oral objetos desde el uno y hasta el que sabe (máximo veinte) siguiendo el orden de la serie numérica	206	98	96	96	98	98
3	Identifica el valor de las monedas	226	95	89	95	95	97
4	Escribe números que le son dictados. (Número menor que tres)	301	89	72	79	91	98
5	Identifica números que sirven para medir longitudes.	320	89	79	87	89	92

Tomada de Backhoff y col. 2008:55.

4.2. Tercero de primaria

La evaluación del aprendizaje en este grado, así como la de sexto de primaria está alineada a los Planes y programas para la educación primaria de la Secretaría de Educación Pública de 1993, que son los vigentes en toda la nación actualmente. Estos planes y programas contemplan seis ejes temáticos, de los cuales cuatro son evaluados por la prueba Excale de tercero: Los números, sus relaciones y sus operaciones, Medición, Geometría y Tratamiento de la información. De los dos restantes, uno no se estudia hasta cuarto grado (Procesos de cambio) y otro incluye para tercero de primaria sólo nociones tan elementales que se decidió no incluirlo sino hasta el Excale de sexto grado (Predicción y azar).

**ILUSTRACIÓN 2. PORCENTAJE DE ALUMNOS POR NIVEL DE LOGRO, POR ESTRATO EDUCATIVO.
EXCALE DE MATEMÁTICAS, TERCERO DE PRIMARIA, APLICACIÓN 2006**



Tomada de Backhoff y col. 2007:39.

En la ilustración 2 puede observarse el porcentaje de alumnos en el país que se encuentra en cada uno de los cuatro niveles de logro. También se muestran los porcentajes en que se encuentran los niños de

cuatro estratos educativos que atienden a porcentajes elevados de la matrícula: Educación indígena, Escuela rural pública, Escuela urbana pública y Escuela Privada⁶.

Como puede verse, en tercero de primaria, el nivel Por debajo del básico es el que agrupa una mayor cantidad de niños (40 %) a nivel nacional. Los alumnos de este nivel "pueden leer y escribir números de tres cifras; calculan sumas sin transformación mediante el algoritmo convencional y resuelven problemas aditivos (sencillos) con números naturales, utilizando procedimientos informales. También identifican objetos cotidianos desde diferentes perspectivas; comparan longitudes y superficies de manera perceptual y utilizando unidades no convencionales, y responden preguntas referidas a la frecuencia a partir de información contenida en tablas" (Backhoff y col., 2007:142).

Por otro lado, estos niños no logran 'identificar el antecesor y el sucesor de un número de tres cifras; ordenar y comparan números de tres cifras; leer y escribir números de cuatro cifras sin cero intermedio; calcular sumas con transformación y restas sin transformación mediante el algoritmo convencional; además no resuelven problemas aditivos hasta de tres cifras cuya solución implica una suma con incógnita en el estado final, ni realizan repartos equitativos con ayuda gráfica, utilizando modelos continuos y discretos. Por otro lado, no identifican figuras básicas (círculos, triángulos, cuadrados y rectángulos) en composiciones; no comparan ni miden longitudes utilizando el centímetro, no resuelven problemas utilizando la hora y el minuto como unidades de tiempo, ni leen o interpretan pictogramas sencillos ni gráficas de barras'⁷ (Backhoff y col., 2007:142). Los niños del nivel Por debajo del básico tampoco cuentan con los conocimientos y habilidades de los niveles Medio y Avanzado⁸.

Los conocimientos y las habilidades de estos niños son todavía incipientes, e insuficientes para afrontar el aprendizaje de la asignatura en grados superiores. En estratos escolares de poblaciones desfavorecidas el porcentaje de niños en esta situación es, lamentablemente, aún mayor: 72 % en Educación indígena y 52 % para el estrato Rural público.

4.3. Sexto de primaria

La evaluación del aprendizaje en este grado está alineada, como se comenta arriba, a los planes y programas de educación básica de la SEP, de 1993, que son los vigentes actualmente en el país. El Excale de sexto de primaria de Matemáticas incluye contenidos de los seis ejes temáticos del currículo.

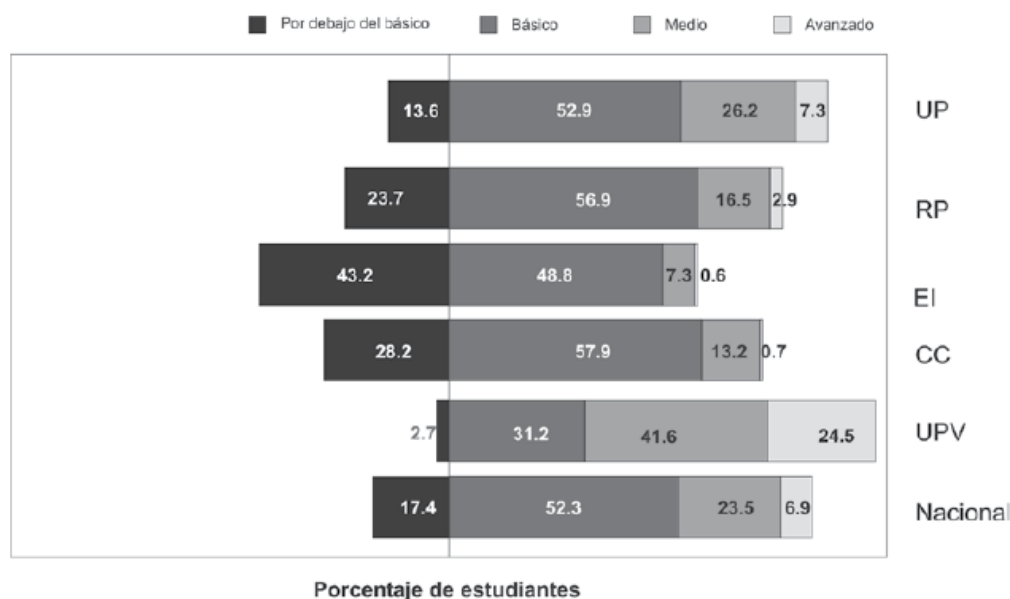
En la ilustración 3 puede observarse el porcentaje de alumnos en el país que se encuentra en cada uno de los cuatro niveles de logro. También se muestran los porcentajes en que se encuentran los niños de cinco estratos educativos que atienden a porcentajes elevados de la matrícula: Escuela urbana pública (UP), Escuela rural pública (RP), Escuelas indígenas (EI), Cursos comunitarios (CC) y Escuela Privada (UPV).

⁶ Se agrega a los estratos educativos presentados en una nota anterior a ésta, la Educación indígena, que se define como "escuelas públicas ubicadas en localidades de población indígena, donde se imparte educación bilingüe bicultural. Los cursos comunitarios no entraron en esta comparación por razones que sería largo explicar, pero que se detallan en el capítulo 6 de Backhoff y col. (2007).

⁷ Se utilizó una comilla simple en lugar de dos comillas para la cita, indicando que el párrafo está parafraseado. Esto es debido a que el original indica lo que sí pueden hacer los niños del nivel Básico, mientras que aquí se está utilizando para describir lo que no pueden hacer los niños del nivel Por debajo del básico. Esto implicó el cambio de algunos verbos y la inclusión de algunas negaciones en el texto.

⁸ Es necesario comentar que además de lo citado (que corresponde al nivel Básico), los niños del nivel previo al básico tampoco pueden hacer lo que se menciona en los niveles Medio y Avanzado, pero nos pareció que podría ser agobiante leer tantas descripciones concatenadas. Se remite al lector interesado a la página 141 de Backhoff y col., 2007.

ILUSTRACIÓN 3. PORCENTAJE DE ESTUDIANTES POR NIVEL DE LOGRO, POR ESTRATO EDUCATIVO. EXCALE DE MATEMÁTICAS, SEXTO DE PRIMARIA, APLICACIÓN 2005.



Tomada de Backhoff y col. 2006:69.

En este grado, a nivel nacional, la mayoría de los niños (52 %) está en el nivel básico. Los niños de este nivel "resuelven problemas con una operación que implique sumas o restas con números de hasta cuatro cifras; además comparan decimales con el mismo número de cifras. Asimismo, calculan el promedio de números naturales en contextos conocidos"; "leen, ordenan y comparan números naturales; además resuelven problemas sencillos con números naturales, decimales y fraccionarios que impliquen una operación en contextos conocidos. Adicionalmente, calculan perímetros y áreas de triángulos y cuadriláteros dentro de una retícula. Asimismo, interpretan información contenida en gráficas y tablas sencillas" (Backhoff y col., 2006:68).

Por otra parte, estos niños no leen, comparan ni ordenan números decimales, ni fraccionarios, ni resuelven con ellos problemas sencillos de suma o resta; además no logran resolver problemas con números naturales que impliquen dos o tres operaciones. No clasifican figuras con base en sus propiedades geométricas; no calculan áreas mediante el uso de fórmulas, ni calculan volúmenes de figuras mediante el conteo de unidades cúbicas; no identifican puntos en croquis, planos y mapas, así como puntos en el primer cuadrante de un plano cartesiano. Asimismo, no interpretan información contenida en gráficas y tablas que contienen datos; no resuelven problemas sencillos de probabilidad que no impliquen realizar análisis combinatorio; ni resuelven problemas de proporcionalidad⁹ (Backhoff y col., 2006:68). Tampoco pueden hacer lo señalado en la misma página para el nivel Avanzado.

Aunque esta descripción nos permite constatar un avance con respecto a lo que puede hacer la mayoría (relativa) de los niños de tercero de primaria, los conocimientos y habilidades descritos son apenas

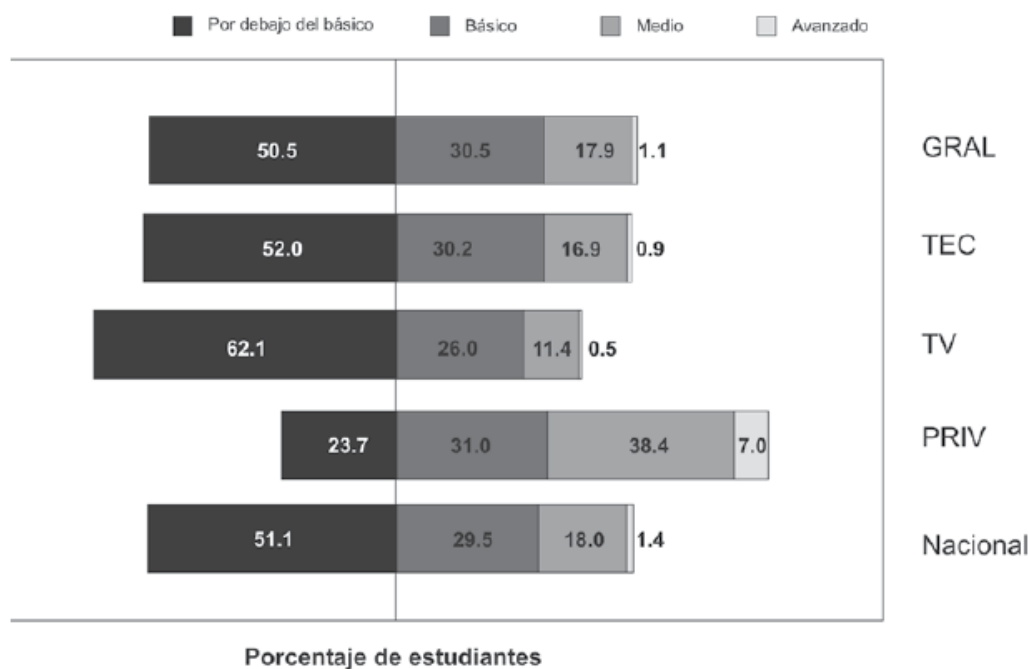
⁹ La cita se hace con una sola comilla en lugar de dos, indicando que se parafraseó el original. En éste, las descripciones estaban dadas en términos positivos, pues explican lo que puede hacer un alumno que se encuentre en el nivel Medio. Sin embargo, para nuestro texto las transformamos en términos negativos, pues se están refiriendo a lo que no puede hacer un alumno de nivel Básico.

suficientes para seguir aprendiendo en secundaria. Además, debe tomarse en cuenta que un 17 % de los niños de sexto no están adquiriendo ni siquiera las habilidades y conocimientos de este nivel. No pretendemos ser pesimistas; por supuesto que también podría ponerse énfasis en los porcentajes de niños que están logrando una cantidad adecuada de conocimientos y habilidades curriculares (el nivel medio) o los que prácticamente adquieren todos los contenidos curriculares (nivel avanzado). Es importante notar que hay buenas proporciones de niños que están llegando a estos niveles, pero para este trabajo aún más importante nos parece hacer notar que hay niños que requieren más apoyo para lograrlo, especialmente en las poblaciones desfavorecidas: en sexto de primaria, un 43 % de niños de Educación indígena y un 28 % de los que están en Cursos comunitarios están en el nivel Por debajo del básico.

4.4. Tercero de secundaria

La evaluación del aprendizaje de este grado está alineada al plan y programas de estudios de secundaria de 1993, puesto que éstos eran vigentes en el momento de la aplicación del Excale de Matemáticas de tercero de secundaria (2005). El currículo de este nivel define cinco áreas temáticas: Aritmética, Álgebra, Geometría, Presentación y tratamiento de la información y Probabilidad; todas ellas fueron evaluadas en la prueba. La distribución de los estudiantes en niveles de logro, tanto nacionalmente como a partir de cuatro estratos educativos¹⁰ se presenta en la Ilustración 4.

ILUSTRACIÓN 4. PORCENTAJE DE ESTUDIANTES POR NIVEL DE LOGRO, POR ESTRATO EDUCATIVO. EXCALE DE MATEMÁTICAS, TERCERO DE SECUNDARIA, APLICACIÓN 2005



Tomada de Backhoff y col. 2006:121.

¹⁰ A diferencia de primaria y preescolar, en secundaria los estratos corresponden con modalidades educativas específicas: Secundarias generales (Gral), Secundarias técnicas (Tec), Telesecundarias (TV) y Secundarias privadas (Priv).

Como puede observarse, a nivel nacional el nivel de logro en que se agrupan más estudiantes es Por debajo del básico. En este nivel los alumnos “resuelven problemas que implican una adición o una sustracción con números naturales, enteros, decimales o fraccionarios, así como problemas de multiplicación y división con números naturales. Así mismo, establecen relaciones entre una tabla de valores y su gráfica, en funciones lineales o cuadráticas. Adicionalmente, identifican figuras o cuerpos geométricos a partir de sus elementos (lados, caras, etcétera). Finalmente, estiman y comparan la probabilidad de eventos simples” (Backhoff y col. 2006:120).

Por otra parte, estos alumnos no ‘resuelven problemas que implican dos o mas operaciones (adición, sustracción, multiplicación, división y potenciación) con números naturales y enteros. No identifican situaciones de proporcionalidad. No Suman ni restan polinomios, ni resuelven ecuaciones de primer grado con una incógnita, así como sistemas de ecuaciones con dos incógnitas. Igualmente, no utilizan las propiedades de las Figuras (por ejemplo: ángulos, lados y diagonales) en la resolución de problemas de medición. Tampoco calculan el perímetro ni el área de Figuras básicas (triángulos, cuadriláteros y polígonos regulares) ni el volumen de cuerpos geométricos¹¹’ (Backhoff y col., 2006:120). Tampoco pueden realizar las tareas descritas para los niveles Medio y Avanzado en la misma página.

Una vez más, aunque la descripción nos permite ver un avance en las habilidades y conocimientos con respecto a los que tenía el grueso de los alumnos de los grados escolares anteriores evaluados, reconocemos claramente que los contenidos que dominan los estudiantes de este grado son insuficientes para seguir aprendiendo matemáticas sin dificultades en los siguientes niveles educativos. Además, esta situación es más grave para los estudiantes de Telesecundaria, puesto que en dicho estrato educativo es de 62 el porcentaje de estudiantes en el nivel Por debajo del básico.

Todo esto sin dejar de reconocer que casi una quinta parte de los estudiantes de tercero de secundaria (unos 370,000 alumnos) alcanzan un buen nivel de dominio de los contenidos matemáticos.

4.5. Análisis a lo largo de la educación básica

Utilizando solamente los datos ya proporcionados sobre los porcentajes de los niños en los diferentes niveles de logro a través de los distintos grados escolares pueden hacerse algunas reflexiones interesantes. En la tabla 5 se presenta un concentrado de los porcentajes de alumnos en cada nivel de logro, por grado escolar, para la población nacional; éstos son tomados directamente de los reportes anuales, puesto que cada grado escolar ha sido evaluado una sola vez a la fecha¹². Junto al grado escolar se muestra, entre paréntesis, el año de medición.

Se puede observar que el mayor porcentaje de estudiantes cambia, entre grados escolares, del nivel Por debajo del básico al Básico. Esto puede deberse al nivel de exigencia de la escala de cada prueba, lo que a su vez se encuentra asociado con el nivel de exigencia del currículo de los distintos grados. Sin embargo, es claro que la mayoría de los alumnos no arriban a los niveles de logro que quisiéramos para todos los mexicanos. La última columna muestra la suma de los dos niveles de logro más bajos. Los datos de ésta nos muestran que a mayor grado escolar, mayor es también el porcentaje de alumnos que se encuentran en los dos niveles de logro más bajos. Por otro lado, puede verse también que para el caso

¹¹ La cita tiene comilla sencilla, en lugar de doble, indicando que el texto original está parafraseado; esto es debido a que en el original se describía en términos positivos lo que pueden hacer los alumnos del nivel Básico, mientras que en nuestro texto se está indicando lo que no pueden hacer los alumnos del nivel Previo al básico.

¹² Exceptuando la réplica que se hizo en 2007 del Excale de sexto de primaria, ya antes referido.

del nivel Avanzado, es decir, aquel en que los niños alcanzan a adquirir prácticamente todo lo que indica el currículo, a mayor grado escolar es menor el porcentaje de alumnos que lo alcanzan.

**TABLA 5. PORCENTAJE DE ALUMNOS POR NIVEL DE LOGRO, POR GRADO ESCOLAR, POBLACIÓN NACIONAL.
EXCALE DE MATEMÁTICAS**

Grado escolar	% Por debajo del básico	% Básico	% Medio	% Avanzado	Suma % Por debajo del básico + Básico
3° Preescolar (aplicado en 2007)	9	49	27	15	58
3° Primaria (aplicado en 2006)	40	20	31	9	60
6° Primaria (aplicado en 2005)	17	52	23	7	69
3° Secundaria (aplicado en 2005)	51	29	18	1	80

Ahora, al construir una tabla similar a la anterior pero referida a las poblaciones más desfavorecidas, de entre las evaluadas en cada aplicación del Excale, encontramos lo siguiente:

TABLA 6. PORCENTAJE DE ALUMNOS POR NIVEL DE LOGRO, POR GRADO ESCOLAR, POBLACIÓN MÁS DESFAVORECIDA EN CADA CASO (SE INDICA JUNTO AL GRADO ESCOLAR). EXCALE DE MATEMÁTICAS.

Grado escolar / población referida	% Por debajo del básico	% Básico	% Medio	% Avanzado	Suma % Por debajo del básico + Básico
3° Preescolar / Cursos comunitarios (3% de la matrícula nacional) (aplicado en 2007)	22	56	16	6	78
3° Primaria / Educación indígena (6 % de la matrícula nacional) (aplicado en 2006)	72	14	12	2	86
6° Primaria / Educación indígena (5 % de la matrícula nacional) (aplicado en 2005)	43	49	7	1	92
3° Secundaria / Telesecundaria (20 % de la matrícula nacional) (aplicado en 2005)	62	26	11	1	88

Puede observarse que la tendencia observada en la tabla de datos nacionales se reproduce aquí, pero con números más alarmantes: a mayor grado educativo, mayor porcentaje de alumnos en los dos niveles más bajos, con excepción del tercero de secundaria en telesecundaria, que tiene un porcentaje cercano al de sexto de primaria en educación indígena, pero ligeramente inferior. Debe considerarse que en esta tabla se está refiriendo la población más desfavorecida de cada grado evaluado, y que éstas no necesariamente son equivalentes para esta comparación. Sin embargo, los datos que señalan la necesidad de establecer políticas que apunten específicamente a mejorar los niveles de logro educativo de las poblaciones más desfavorecidas son elocuentes. El Sistema Educativo Mexicano debe plantearse seriamente el reto de comprender a fondo las razones de estas tendencias, para aspirar a revertirlas.

A partir de los datos presentados en este trabajo, podemos plantear algunas hipótesis que buscan explicar la disminución del porcentaje de estudiantes en los niveles superiores al aumentar el grado escolar:

- Los alumnos de un grado bajo (pongamos, tercero de preescolar) que se encuentran en los niveles Por debajo del básico y Básico permanecen en general en esos mismos niveles al arribar a posteriores grados escolares, dado que la cantidad de cosas que han aprendido no les da elementos suficientes para remontar su historia con respecto a la asignatura y arribar a mejores comprensiones y adquisiciones de conocimientos y habilidades en grados posteriores.

Posiblemente los alumnos que estén en esta situación y participen en programas remediales muy bien enfocados podrán mejorar significativamente su desempeño en grados posteriores, pero no lo harán, en general, sin ayuda externa.

- b. Algunos alumnos que se encontraban en los niveles medio y avanzado pueden tener dificultades en su historia personal, familiar, escolar, social, desde enfermedades hasta participación (como víctima o atacante) en actos violentos, desde ser trabajadores remunerados hasta conflictos familiares, que llevan a que la calidad de su educación disminuya localizándole, al subir de grado, en niveles de logro inferiores.
- c. Los currículos de grados superiores incluyen una complejidad que crece a un ritmo mayor que la capacidad de los estudiantes por retener las cantidades de información y de adquirir las habilidades que se les requieren, de manera que, otra vez, algunos alumnos que se encontraban en los niveles medio y avanzado bajan a los niveles inferiores.
- d. Desde los primeros grados escolares se muestran brechas educativas entre los niños que tienen mejores condiciones socioeconómicas y los que las tienen peores. Al ascender en grados escolares, las familias en condiciones socioeconómicas pobres tienen mayores dificultades para apoyar correctamente el aprendizaje de los niños (por ejemplo, aumenta la dificultad de las preguntas que deben contestarse en una tarea, o es de mayor costo el material que se pide para una clase, o la escuela secundaria está en general más lejos que la primaria, en el caso de zonas rurales), de modo que una vez más, estudiantes que se encontraban en los dos niveles altos, bajan.

De ser ciertas estas hipótesis, debe notarse que las ocasiones para bajar a los dos niveles de logro inferiores son más que las de mejora, de tal suerte que, metafóricamente, se conforma una especie de agujero negro de la educación, que tira hacia abajo cuando no hay una oportunidad de acceso real al conocimiento, o cuando ésta es desigual.

5. CONCLUSIONES

Las pruebas Excale, que cuentan ya con resultados para cada uno de los grados de educación básica para los cuales están previstas, permiten hacer análisis sobre los niveles de logro de dicho nivel educativo en Matemáticas, ofreciendo un panorama sobre el aprendizaje de esta disciplina en México.

Como puede observarse en el plan de pruebas, los Excale tienen la intención de ofrecer comparaciones en el tiempo que puedan mostrar tendencias del aprendizaje en un mismo grado a lo largo de los años, en varios grados a partir de un corte transversal (como hace este trabajo) o incluso siguiendo cohortes de estudiantes. Los Excale han tenido una breve vida todavía, y por ello no es posible mostrar series de tiempo aún. Sin embargo, una intención fundamental de estas pruebas es llegar a ello, con el rigor técnico que permita obtener conclusiones sólidas sobre la evolución del logro educativo en México.

Uno de los elementos que brindan las pruebas Excale y que podrían aumentar nuestra comprensión de las deficiencias matemáticas de los estudiantes de educación básica son los datos en que se indican los porcentajes de acierto de los estudiantes por cada contenido evaluado. Dado que los contenidos evaluados se estructuran en líneas de evaluación, su análisis por parte de especialistas en didáctica de las

matemáticas podría mostrar elementos para mejorar las propuestas curriculares, los materiales didácticos, los cursos de actualización y los programas específicos de mejora educativa.

Encontramos que la mayor parte de los estudiantes en el país no alcanzan, al finalizar los grados escolares evaluados, un dominio adecuado de los contenidos del currículo nacional. Es posible, a partir de la información completa de los reportes de resultados de los Excale, analizar a profundidad los temas en los que los estudiantes tienen menor dominio, posibilitando acciones focalizadas a reforzar la enseñanza y el aprendizaje de éstos; pero más allá de lo escolar, sabemos que diversos factores extraescolares, sociales y familiares influyen en los resultados. En el presente trabajo no se profundizó a este respecto, aunque algunos indicios se pueden encontrar en el hecho de que, para el caso de las primarias, los estratos Escuela pública urbana y Escuela pública rural son en realidad una misma modalidad educativa, en dos tipos de localidades diferentes, y a pesar de ello, muestran resultados de logro escolar muy distintos a lo largo de los grados escolares evaluados.

Conocer los niveles de logro de los estudiantes del país es sólo un primer paso para trabajar por la mejora de la educación. La situación de la educación matemática básica presenta importantes retos para el Sistema Educativo Mexicano, especialmente en las poblaciones socialmente desfavorecidas. Resulta útil recalcar aquí que varios estudios cuyos resultados no ha sido posible presentar en este trabajo (Backhoff y col. 2007b, 2007c) muestran cómo diversas variables de contexto, incluyendo algunas que no tienen relación con la escuela, influyen poderosamente y de manera decisiva en los resultados educativos. Resulta así fundamental considerar políticas integrales que afronten simultáneamente diversos elementos de la realidad escolar y contextual de los estudiantes para poder lograr cambios significativos.

Trazamos en los resultados una tendencia a que haya más estudiantes en los niveles de logro más bajos conforme se avanza en grados escolares. Esto puede deberse fundamentalmente a que quien se encuentra en una situación de no haber adquirido los suficientes aprendizajes en un grado escolar, difícilmente podrá aprovechar correctamente las clases en grados posteriores, creándose una situación de aprendizajes deficientes que sólo podría resolverse incidiendo directamente sobre ella, con la intención específica de remontar el deterioro del logro. Las dificultades personales, familiares y escolares, la complejidad creciente de los currículos (cuando no es acorde con la adquisición creciente de conocimientos y habilidades) y las mayores dificultades de las familias y comunidades educativas con bajos niveles socioeconómicos para apoyar correctamente el aprendizaje escolar son otros elementos que podrían estar tirando hacia abajo el nivel de logro de los estudiantes.

Sería necesario llevar a cabo investigación focalizada en estos elementos para constatar si actúan de la manera en que se indica. Para ello tendrían que llevarse a cabo estudios de factores asociados al aprendizaje y de seguimiento de políticas educativas, basados en evaluaciones de gran escala; pero también una gran cantidad de estudios puntuales, con pequeñas poblaciones y de corte más cualitativo, que den cuenta de los elementos más sutiles de la relación que guarda el logro escolar con las circunstancias personales, familiares y escolares de los alumnos. Los meta análisis que sintetizen las experiencias de diferentes grupos de investigación con este interés serían una vía importante para comprender el deterioro del logro y tomar decisiones de política educativa y de gestión de programas remediales mejor fundamentados.

Es la visión de quienes trabajamos en estas pruebas, que cada vez más profesores, investigadores, autoridades educativas y grupos ciudadanos con interés en la educación miren la buena cantidad de información que se ofrece a partir de los resultados de los Excale para guiar y retroalimentar sus

decisiones, de manera que la evaluación sistemática y oportuna se convierta finalmente en un elemento constante del diseño curricular, el análisis y la implementación de políticas educativas nacionales, estatales, regionales y sectoriales, la reflexión informada de la ciudadanía y la discusión colegiada de los especialistas. Creemos que sólo entonces la evaluación educativa de gran escala estará justificando su existencia.

REFERENCIAS BIBLIOGRÁFICAS

- Backhoff, E.; Monroy, L.; Peon, M.; Sánchez Moguel, A.; y Tanamachi M. L. (2005). *Exámenes de la Calidad y el Logro Educativos (I aprendizaje del Español y las Matemáticas en la educación básica en México. Sexto de primaria y tercero de secundaria*. México: INEE.
- Backhoff, E.; Andrade, E.; Sánchez Moguel, A.; Peon, M.; Bouzas, A. (2006). *El aprendizaje del Español y las Matemáticas en la educación básica en México. Sexto de primaria y tercero de secundaria*. México: INEE.
- Backhoff, E.; Andrade, E.; Sánchez Moguel, A.; Peon, M. (2007). *El aprendizaje en tercero de primaria en México. Español, Matemáticas, Ciencias naturales y Ciencias sociales*. México: INEE.
- Backhoff, E.; Bouzas, A.; Hernández, E.; García, M. (2007b). *Aprendizaje y desigualdad social en México. Implicaciones de política educativa en el nivel básico*. México: INEE.
- Backhoff, E.; Bouzas, A.; Contreras, C.; Hernández, E.; García, M. (2007c). *Factores escolares y aprendizaje en México. El caso de la educación básica*. México: INEE.
- Backhoff, E.; Andrade, E.; Sánchez Moguel, A.; Peon, M. (2008). *El aprendizaje en tercero de preescolar en México. Lenguaje y comunicación. Pensamiento matemático*. México: INEE.
- Backhoff, E.; Andrade, E.; Sánchez Moguel, A.; Peon, M. (2008). *Estudio comparativo del aprendizaje en sexto de primaria en México 2005 – 2007: Español y Matemáticas*. México: INEE.
- Instituto Nacional para la Evaluación de la Educación (2005a). *¿Qué es el INEE? Sus fines y estrategias al servicio de la sociedad*. México: INEE.
- Instituto Nacional para la Evaluación de la Educación (2005b). *Excale. Exámenes de la Calidad y el Logro Educativos. Proceso de construcción y características básicas*. México: INEE.
- Instituto Nacional para la Evaluación de la Educación (2006). *Plan maestro de desarrollo 2007-2014*. México: INEE.
- Nitko, A. (1994). A model for developing curriculum-driven criterion-referenced and norm-referenced national examinations for certification and selection of students. Ponencia presentada en la *Conferencia Internacional sobre Evaluación y Medición Educativas*, de la Asociación para el Estudio de la Evaluación Educativa en Sudáfrica (ASSESA).

LAS PRÁCTICAS DE LOS COMITÉS DE EVALUACIÓN DEL APRENDIZAJE ESCOLAR. ALGUNAS PARADOJAS A PARTIR DEL CASO MEXICANO

Adriana Andrade

Revista Iberoamericana de Evaluación Educativa 2009 - Volumen 2, Número 1

<http://www.rinace.net/riee/numeros/vol2-num1/art11.pdf>

Fecha de recepción: 15 de marzo de 2009

Fecha de comunicación de dictamen: 14 de abril de 2009

Fecha de aceptación: 18 de abril de 2009

Los comités de evaluación del aprendizaje escolar pueden ser estudiados como un fenómeno sociocultural en el cual las personas involucradas interactúan entre sí en determinados espacios y momentos, dentro de un marco de consideraciones de la evaluación como una cuestión técnica con connotaciones de forma y contenido permisibles. Este marco de consideraciones sin embargo, no es un monolito sino que en las prácticas cotidianas de evaluación se pueden identificar tanto los mecanismos de control, como puntos de ruptura y oportunidades que fomentan la producción de significados plurales de la evaluación, por un lado, y la posibilidad de ampliar la cultura ciudadana desarrollando competencias de información, de investigación, y de diálogo que contribuye a la resolución no violenta de conflictos en el campo educativo.

En este trabajo argumentaré que en medio de las tensiones sociales que se viven en torno a la evaluación de la calidad de la educación en México¹ la comprensión de los comités de evaluación educativa como fenómeno sociocultural permite:

- explorar y alimentar propuestas de aproximaciones particulares hacia la mejora de la educación, que hacen de las acciones y los datos, instrumentos tangibles que posibilitan la comprensión de la realidad, sin caer en dogmatismos totalizadores ni simplistas;
- abrir opciones de gestión educativa como un asunto de interés público, no como un hecho dado, sino como un proceso que cambia y se intensifica, a lo largo de un proceso de negociación de sus significados sociales;
- nutrir un horizonte ético para crear influencia y establecer relaciones de equidad en el campo educativo.

Para fundamentar esta idea presento el caso de los comités del Instituto Nacional de Evaluación Educativa (INEE)² de México, como una práctica contextualizada que entrelaza la participación social y la rendición de cuentas en el campo de la educación, como una política pública que cobra vida en acciones que socializan el conocimiento generado, que construyen valores comunes, que asumen compromisos y acuerdos con los demás, que fomentan relaciones de complementariedad y de colaboración, dentro de ciertos límites y posibilidades.

En resumen se trata de una práctica social compleja que presenta paradojas que influyen en la reconfiguración del espacio público de la educación, sus límites y posibilidades, con elementos de uno de los diversos aspectos de la problemática de la calidad de la educación: el desarrollo de competencias ciudadanas para la participación social y la rendición de cuentas inteligente y comprometida.

El artículo se estructura en seis partes. La primera parte presenta información básica general sobre la fecha de creación del INEE y su función social, y sobre la conformación de cada uno de los comités que operan en el Instituto. La segunda parte trata sobre la situación de comunicación de los comités, en él se describen las posiciones que ocupan los participantes en los comités y las características de sus

¹ A nivel macrosocial, la Alianza por la Calidad de la Educación que promueve el gobierno federal mexicano se actualiza en un plan de reforma educativa que no ha logrado un consenso con todos los agentes sociales vinculados en el proceso educativo sobre una diversidad de aspectos, entre ellos la evaluación externa y uniforme del aprendizaje de los estudiantes.

A nivel microsocio, el desconocimiento de los límites y posibilidades de la evaluación de la calidad educativa lleva a utilizarla en forma inadecuada en muchas escuelas, basados en resultados aislados sin considerar el contexto.

² Organismo público descentralizado, de carácter técnico, con personalidad jurídica y patrimonio propio.

interrelaciones apoyadas en las rutinas institucionalizadas, sus contenidos y formas, así como los elementos emergentes por el carácter singular de cada grupo.

En los apartados siguientes se habla de los discursos, los objetos y prácticas de los comités de evaluación que en conjunto conforman un lenguaje de contención que opera como marco de referencia para vivir a través de y actuar hacia los ordenamientos sociales del campo de la educación.

En la última parte se sistematizan las paradojas encontradas, y se puntualiza cómo en el estudio de caso del INEE, las prácticas de los comités de evaluación como fenómeno sociocultural de participación social y rendición de cuentas, contribuyen a la reconfiguración del espacio educativo como una cuestión de interés público.

1. LA ESTRUCTURA Y FUNCIONAMIENTO DE LOS COMITÉS DEL INEE

El Instituto Nacional de Evaluación Educativa (INEE) fue creado en México por decreto Presidencial publicado en el Diario Oficial de la Federación el 8 de agosto de 2002 para ofrecer a las autoridades educativas y al sector privado herramientas idóneas para la evaluación de la educación básica (preescolar, primaria y secundaria) y media superior. El propósito es conocer el rendimiento académico de los estudiantes a nivel estatal y nacional, y los factores de contexto que lo explican.

Una de sus funciones es diseñar pruebas nacionales alineadas al currículo de las asignaturas de matemáticas, español, ciencias sociales, ciencias naturales, a través de un proceso donde el aspecto central es el trabajo colegiado en el que participan diversos especialistas, grupos de asesores, comités de expertos y personal técnico de la Dirección de Pruebas y Medición del INEE.

Son cinco tipos de comités que intervienen en el proceso de diseño, implementación y apreciación de los resultados obtenidos con los Exámenes de la Calidad y el Logro Educativo (Excale) a nivel nacional:

- Comités académicos,
- Comités elaboradores de especificaciones de reactivos,
- Comités constructores de reactivos,
- Comités de validación y sesgo y
- Comités de niveles de logro.

Hay un Comité Académico de cada prueba: matemáticas, español, civismo, ciencias naturales, y grado correspondiente³, el cual analiza el currículo y elabora una retícula de la asignatura con la estructura de los contenidos curriculares y la importancia relativa de cada uno de ellos. Este comité integra una tabla de contenidos de la prueba y explicita las áreas temáticas y habilidades intelectuales a evaluar, así como el enfoque de evaluación.

El Comité de Elaboración de Especificaciones de Reactivos de cada prueba conoce a profundidad el currículo de la asignatura a la que se dirige el Excale correspondiente y el dominio a evaluarse (conocimientos, habilidades, competencias). Debe precisar los aspectos teóricos y metodológicos del

³ En el Plan General de Evaluación del INEE 2005-2016 se contemplan los siguientes grados escolares: 3º de Preescolar (2007 y 2011), 3º de Primaria (2006, 2010 y 2014), 6º de Primaria (2005, 2007, 2009, 2011, 2013, 2015), 3º de Secundaria (2005, 2008, 2010, 2012, 2014, 2016). En el paréntesis aparecen los años escolares a los que se aplicarán los Excale.

contenido a evaluar, su ubicación en el currículum, su importancia, la habilidad intelectual requerida al alumno, así como las características específicas que debe tener cada reactivo.

Los miembros del Comité de Constructores de Reactivos formulan de manera independiente tres reactivos para cada especificación de un contenido a evaluar, con base en las especificaciones y la normatividad del INEE. Estos reactivos son probados en una muestra pequeña de estudiantes y son revisados por otros especialistas para asegurar que cada uno de los reactivos represente la parte del currículum para la cual se especificó, es decir que tenga congruencia con la especificación, que el lenguaje que utiliza sea adecuado y que el formato sea pertinente (opción múltiple, respuesta corta, ensayo, entre otros ejemplos de formatos).

En el Comité de Validez y Sesgo, de manera independiente cuando menos dos profesores revisan cada reactivo en relación con su contenido y sesgos culturales y de género; y proponen recomendaciones para mejorar los reactivos o justifican su eliminación. En él participan docentes en ejercicio de las 32 entidades federativas del país, que representan a los diversos estratos y modalidades educativas.

El Comité de niveles de logro participa en el análisis e interpretación de resultados de los Excale asociados a variables de contexto, y establece los niveles de competencia alcanzados: por abajo del básico, básico, medio, avanzado. No se tienen puntuaciones por alumno ni por escuela, se conocen los contenidos que dominan las y los estudiantes y los que no.

Uno de los aportes más importantes del trabajo conjunto que realizan los comités de evaluación es establecer una línea base de medición y puntos consecutivos de comparación en el tiempo. Estas cuestiones son estratégicas para que los procesos de evaluación sean válidos, confiables, consistentes y coherentes.

Una característica de la estructura de los comités es la necesidad de multidisciplinariedad, que se logra con la contratación de los servicios de expertos externos de disciplinas específicas afines a los contenidos a evaluar (matemáticas, español, civismo, ciencias naturales) y de otras disciplinas ligadas a las ciencias de la educación y a los tópicos temáticos a tratar (psicólogos educativos, sociólogos, pedagogos, historiadores, filósofos, biólogos, licenciados en educación básica, entre otros); y la supervisión del desempeño bajo el criterio de calidad, oportunidad y costo adecuado. La participación en los comités tiene una visibilidad social que contribuye a dar legitimidad a las acciones que realiza: se acredita la personalidad de los participantes, se firma un contrato de trabajo por obra determinada con honorarios profesionales, se extiende una constancia de participación.

Los comités están conformados por un grupo promedio de 10 personas invitadas por el INEE que participan por proyectos según el tipo de comité de que se trate. El proyecto y la vigencia de un comité inicia en una reunión presencial donde los participantes se presentan, comunican su escolaridad, experiencia en el campo de la educación, interés en participar; se desarrolla de acuerdo a un plan del trabajo encomendado en espacios y tiempos presenciales y/o virtuales; y termina el proyecto y el comité al entregar los resultados esperados.

Entre las características del perfil de los integrantes de los comités están las siguientes: algunas personas colaboran con instituciones de educación superior e investigación, otras con instituciones gubernamentales del campo educativo, o son maestras y maestros en funciones dentro del sistema educativo nacional, algunas personas son autores de libros de texto, hay quienes forman parte de una organización de la sociedad civil orientada a la educación. La escolaridad mínima de los miembros de un

comité es licenciatura, aunque en promedio la mayoría de los integrantes tienen estudios de posgrado. Dentro de los comités participa también personal especializado del INEE y actúa como moderador y mediador de la tarea de los comités.

Los lineamientos establecidos por el INEE, concretados en planes de trabajo de los proyectos encomendados a cada comité, determinan de manera precisa los momentos y espacios donde intervienen los comités, la metodología que deben seguir para desempeñar su función y los formatos donde deben sistematizar los resultados de su trabajo. En conjunto estos elementos de organización contribuyen a darle estructura al grupo de trabajo teniendo como referente principal la tarea encomendada al comité y la orientación a resultados.

Las características del proceso de trabajo de los comités exigen la aplicación de competencias de información, de investigación y de diálogo. En el Cuadro 1. *Competencias de los comités de evaluación*, se resumen las categorías e indicadores de los conocimientos, habilidades y actitudes puestos en escena durante el tiempo que se desempeñan los miembros de un comité en los procesos de diseño, construcción y validación de los Excale. Estas competencias se relacionan con comportamientos, habilidades y actitudes propios de un ámbito de trabajo colegiado y de los aspectos técnicos directamente relacionados con la tarea bajo su responsabilidad.

CUADRO 1. COMPETENCIAS DE LOS COMITÉS DE EVALUACIÓN

COMPETENCIA GENÉRICA	DESCRIPCIÓN
Capacidad de análisis descriptivo	Poder describir el objeto y proceso de una evaluación, cuáles son sus límites y cuales sus características principales
Capacidad de análisis relacional	Poder identificar los factores del entorno que influyen sobre el objeto y proceso de una evaluación.
Capacidad de administración de información	Poder utilizar la información existente para tomar decisiones sobre el enfoque de evaluación más conveniente.
Capacidad de administración de información	Poder identificar las modalidades de relevamiento, procesamiento, análisis y sistematización de información.
Capacidad de administración de información	Poder anticipar los requerimientos de información, saber presentar los resultados para que cumplan su función.
Capacidad de análisis estratégico	Poder mantener el enfoque, fijar prioridades y justificarlas.
Capacidad de argumentación	Poder argumentar en forma sistemática y fundamentada la postura individual.
Capacidad de orientación a resultados	Poder administrar el proceso de evaluación, distribución eficiente de tiempos y recursos, y un cumplimiento eficaz de los requerimientos.
Capacidad de juicio moral	Poder obrar éticamente, con prudencia, con justicia.
Capacidad de autorregulación	Poder trabajar en forma autónoma, ser flexible y ser autocrítico.
Capacidad de autorregulación	Poder aprender
Capacidad de trabajar en equipo	Poder trabajar en equipo, valorar la diversidad cultural, tener responsabilidad social.
Capacidad de comunicación oral y escrita	Poder expresarse con claridad
Habilidades interpersonales	Poder ser empático con las otras personas miembros del comité
Capacidad de negociación y resolución de problemas	Poder encontrar soluciones, transferir experiencias a nuevas situaciones, tolerar imprevistos
Capacidad de negociación y resolución de problemas	Poder construir, mediante el diálogo acuerdos aceptables para todos a partir de intereses diferentes, parecidos o iguales.
Conocimientos sobre el área de estudio y procedimientos	Dominio experto de contenidos, procedimientos y productos de la tarea; saber dónde buscar información; experiencia en un tema en particular.
Capacidad de investigación	Poder objetivar y ampliar la reflexividad sobre el campo de estudio.
Compromiso ético y compromiso con la calidad	Poder convivir y hacerlo en forma comprometida con valores universales, derechos humanos, sociales, individuales, y códigos profesionales.
Visión de futuro e innovación	Poder pensar de manera prospectiva y proponer formas nuevas de incidir en la realidad.

FUENTE: elaborado por la autora con base en las notas del Diario de Campo

Al interrelacionarse los miembros de un comité, su desempeño sirve como referente para ajustar los comportamientos propios y de los otros, y para mejorar el nivel de competencia grupal. La precisión sobre el producto esperado a ser construido con los otros y el procedimiento para llevar a cabo la tarea son factores que influyen de manera positiva en la interacción, contribuyen a mantener: una administración eficiente del tiempo, la claridad del enfoque de la tarea que los ocupa, una orientación a resultados válidos y confiables. La disposición de los participantes para hacer un alto y revisar y especificar sus puntos de vista motiva un clima de intercambio abierto de información. Hay situaciones en las que los puntos de vista difieren e incluso pueden llegar a oponerse, pero la responsabilidad de entregar los resultados comprometidos fomenta una disposición a escuchar y establecer acuerdos que permitan cumplir el objetivo común.

El resultado obtenido en colectivo satisface las expectativas instituidas, como mínimos indispensables asegurados por la mediación del representante del INEE. La mediación se orienta hacia el logro de una mayor comunicación e interacción, desde las relaciones (presenciales o virtuales) entre los actores participantes y desde los procesos de producción de los productos esperados (textos impresos, digitalizados). Las reglas establecidas por sí mismas no pueden generar los productos de la evaluación, sino los procesos de los comités de evaluación, en los cuales pueden surgir elementos que irrumpen lo establecido y lo transforman agregando valor con modificaciones o incluso innovaciones. De estos elementos "instituyentes" se trata el siguiente apartado sobre la situación de comunicación de los comités.

2. LA SITUACIÓN DE COMUNICACIÓN DE LOS COMITÉS

Para entender los rasgos relevantes de la situación de comunicación de los comités de evaluación en el caso de estudio hay que remitirse a las posiciones que ocupan cada uno de los participantes y las características de las interrelaciones que establecen.

En cuanto a las posiciones que ocupan las personas miembros del comité, todas tienen voz y voto en las decisiones que se toman para generar los resultados esperados. La autoridad relativa de cada persona está respaldada en principio por el reconocimiento de su capacidad profesional al haber sido invitado al comité, por la narrativa de su trayectoria con el que se presenta en la fase introductoria de los trabajos del comité, pero sobre todo por el fondo y la forma de argumentar durante las intervenciones, de carácter libre, durante las reuniones colectivas. En general el propio grupo se autorregula para que se respeten los turnos para tomar la palabra, o no alargarse mucho en una argumentación; en caso de que no se llegue a acuerdo entre todos los miembros del comité, el representante del INEE en su calidad de mediador ofrece opciones para buscar el consenso, por ejemplo a través de una votación.

El ambiente de la situación de comunicación inicial es adecuado para orientarse hacia la tarea y contener las relaciones interpersonales, los miembros del comité se relacionan con los demás sin estar sometido a presiones, es un tiempo reconocido como valioso, y su desempeño está remunerado. La dinámica del ambiente de la situación de comunicación es cambiante en función de la reacción de cada persona frente a lo que hacen los demás. Si la tarea se enfrenta a obstáculos por falta de información el mediador interviene para prevenir que se pierda el sentido de lo que están haciendo; esto se logra procurando acceder a la información necesaria de manera oportuna, como parte de la capacidad de planeación y organización del Instituto. En el caso de observar un ambiente turbulento por el enfrentamiento de

posiciones opuestas entre los participantes al elegir vías de acción contrapuestas para conseguir los resultados buscados, es muy importante la flexibilidad y versatilidad de los participantes para identificar recursos y opciones capaces de cambiar el rumbo de los acontecimientos, disminuir la tensión en las relaciones interpersonales, centrarse en la tarea y alcanzar los resultados esperados.

En este aspecto otro factor decisivo es la motivación hacia el logro, actitud orientada a la resolución no violenta de conflictos y el nivel de aspiración a demostrar ser profesionales de valía para la labor encomendada que cada participante tiene. Esta motivación está relacionada por un lado con el reconocimiento de capacidad técnica que se obtiene al incluir en el currículo profesional la participación en los comités del INEE, y por otro lado, la posibilidad de ser invitado para formar parte de otros comités, de ejercer influencia para intensificar el interés público en la educación, de obtener una remuneración por los servicios profesionales prestados.

En cuanto a las características de las interrelaciones que establecen los miembros de un comité, durante las reuniones de los comités cada persona se comunica con los demás a partir de los referentes preexistentes individuales, de acuerdo a su formación disciplinar, su trayectoria profesional, su carácter, su ciclo de vida, su género, entre otros factores. Estos referentes configuran los mapas asociativos de cada participante, mismos que van compartiendo a lo largo de las jornadas del comité que puede durar entre 3 y 5 días de trabajo presencial en las instalaciones preparadas por el Instituto (salas de juntas internas o extramuros) y períodos variables de trabajo individual virtual en tiempos y espacios diferenciales ya que pueden trabajar fuera de las instalaciones del Instituto en horarios variables y comunicarse a través del Sistema Integral de Administración de Excale – SIAE que opera como administrador de documentos que hayan sido redactados, corregidos o revisados por los miembros de los comités y que deben ser enviados al INEE y recibidos del INEE únicamente por medio del sistema, lo que contribuye a asegurar que el trabajo se mantiene organizado, que está documentado el proceso de trabajo con las versiones actualizadas de los formatos llenados con los resultados individuales y colectivos de los comités.

Para que los miembros del comité puedan interconectar los mapas asociativos se requiere de un acoplamiento entre ellos, el cual se construye sobre las bases de rutinas institucionalizadas a través de un conjunto de actos y comportamientos estandarizados en la situación de comunicación de los comités, que facilitan el encuentro entre personas distintas, sin que medien grandes esfuerzos, adquiere cierto carácter ritual y se sostiene de manera durable con el apoyo del aparato institucional: agenda, horarios, documentos, servicios.

En el caso del INEE la acción de evaluación educativa comprende un proceso con una fisonomía administrativa característica: se planea el diseño, la construcción, la aplicación y la validación de los Excale. Cada una de las fases de este proceso considera el uso racional de recursos humanos, financieros y materiales; se supervisan las acciones para que se cumplan de acuerdo a lo estipulado; se estimula la participación de quienes han sido convocados y se justifica toda diferencia entre lo planeado y lo alcanzado.

La documentación del proceso del trabajo de los comités es un apoyo básico para verificar que se cuenta con los insumos, con los resultados, las interrelaciones entre acciones, la distribución de espacio y tiempo. En el Cuadro 2 las columnas señalan las categorías del proceso de diseño, construcción y validación de los Excale, cuándo, quiénes, cómo, qué; en las filas se especifican las fases de la administración del proceso, desde la planeación hasta la validación. La visualización del proceso en esta

tabla es un instrumento que facilita la ubicación y colaboración entre los miembros de los distintos comités, y permite una comprensión general del sentido del mismo.

CUADRO 2. PROCESO DE DISEÑO, CONSTRUCCIÓN Y VALIDACIÓN DE LOS EXÁMENES DE CALIDAD Y EL LOGRO EDUCATIVOS EXCALE DEL INEE

Fases	Etapas	Participantes	Procedimientos	Productos
I. Planeación General	1.Diseño del Plan General de Evaluación			
	2.Diseño y elaboración de cuestionarios de contexto			
	3.Diseño y desarrollo del sistema informático			
II. Estructuración	4.Diseño de las pruebas	Comité académico	Documentación, capacitación, preparación de materiales, trabajo colegiado	Manual técnico de diseño de los Excale. Réticula curricular de cada examen. Tabla de contenidos de cada examen.
	5.Especificaciones de reactivos	Comité elaborador de especificaciones	Documentación, capacitación, preparación de materiales, trabajo colegiado	Manual técnico para la elaboración de especificaciones. Especificaciones de reactivos de cada examen. Revisiones.
III. Construcción de reactivos	6.Elaboración de reactivos	Comité constructor de reactivos	Documentación, capacitación, trabajo individualy colegiado	Manual técnico para la elaboración de reactivos. Tres reactivos por especificación. Revisiones.
	7.Validación de reactivo	Comité de validación sesgo	Documentación, capacitación, trabajo individualy colegiado	Manual técnico para la validación de reactivos. Reportes de validación.
	8.Piloteo de reactivos y cuestionarios de contexto			
IV. Conformación	9.Selección de reactivos e integración de bloques y formas			
V. Aplicación	11.Muestreo poblacional			
VI. Análisis e interpretación de los resultados	13.Análisis iniciales de resultados			
	14.Establecimiento de niveles de logro	Comité de niveles de logro	Documentación, capacitación, trabajo colegiado	Manual técnico sobre establecimiento de niveles de logro y documento donde se consigne los niveles de logro y puntos de corte de cada Excale.
	15.Elaboración de informe técnico de resultados de aprendizaje			
VII. Validación	16.Estudios de validez de procesos y resultados de las pruebas			

FUENTE: Resumen de la Tabla del mismo nombre del Manual técnico de especificaciones de reactivos de marzo 2008, elaborado por la autora.

En el estudio de caso los involucrados experimentan un tipo de metacognición en la medida que toman conciencia de los lineamientos, estándares, metodología, los internalizan y amplían la comprensión de los factores que explican los resultados de su actividad, sean positivos o negativos. En este proceso de hacer conciencia, se exhibe la capacidad de autorregulación y control de las actividades de evaluación, desde la planeación, la puesta en marcha, hasta el seguimiento y control del proceso de trabajo del comité en su conjunto.

El resultado de este proceso de metacognición es una transformación de un *habitus*⁴, es decir desarrollan un conjunto de esquemas de percepción, de comprensión, de decisión y de acción durables.

Las rutinas tienen un efecto positivo porque se asocian a comportamientos saludables dada su fundamentación en investigaciones científicas, porque contribuyen a elevar la eficiencia o desempeño personal, por estar vinculadas a normas de organización individual, adecuado empleo y distribución del tiempo, métodos de planificación y control de metas y actividades, etc. en la realización de una tarea que presenta desafíos por su complejidad y dinamismo. El conjunto de elementos influye para que se propicie el apego a las rutinas.

Sin embargo, un segundo efecto del proceso de metacognición se refiere a la capacidad potencial de que los miembros del comité puedan seleccionar, valorar o intersectar componentes de formas simbólicas de otros contextos independientes a la evaluación educativa, que reflejen la experiencia acumulada y estructurada y las visiones del mundo que cada uno tiene. A pesar de los beneficios que las rutinas institucionalizadas muestran, hay desventajas, en tanto pueden disminuir el interés, opacar la creatividad e impedir intentar estrategias, emplear métodos o abordajes novedosos y más productivos en determinadas áreas o actividades.

Bajo esta perspectiva en la situación de comunicación de los comités se da lugar a adaptaciones contingentes de las buenas prácticas transmitidas o incluso a innovaciones que pueden generar significados emergentes.

Un ejemplo es la relevancia de la perspectiva de equidad de género que tiende a aumentar en la medida que los participantes integran en su mapa asociativo los derechos humanos, la perspectiva de equidad, la consideración de la población vulnerable. Se observa entonces que en las prácticas de los comités hay la reiteración de una política de comunicación que impida la masculinización de los enunciados, lo cual se refleja en la documentación generada por el comité en cuestión, donde la narrativa subraya el uso de artículos femeninos y masculinos como producto de un acuerdo grupal. Vgr. "las y los estudiantes". Esto no era parte de los lineamientos o metodología establecida institucionalmente, pero termina por modelar el comportamiento del grupo, e incluso con la posibilidad de extenderse a otros grupos que se expongan al contacto con este tipo de acuerdos grupales.

En la medida que los puntos de conexión entre los agentes sociales en situación de comunicación y los mapas cognitivos asociativos de las prácticas de evaluación se expanden en escala, densidad y complejidad -ya que el diseño de los Excale es a nivel nacional, cubre diferentes dominios de conocimiento y habilidades- la información acumulada sobre la evaluación educativa se vuelve una

⁴ Se retoma a Bourdieu y su concepto de *habitus* o de sistema de disposiciones, que se propone sean durables y transponibles; las cuales han sido socialmente construidas a título de estructuras estructurantes y estructuradas, que constituyen el principio de generación y estructuración de prácticas [Bourdieu, 1972]

cuestión abierta a toda la red de relaciones que tejen los miembros de los comités hacia fuera de la institución.

El esquema de percepción que se construye conforma un juicio sobre el objetivo de la educación y sobre la manera que creen es más conveniente para cumplir tal objetivo. Los mapas asociativos compartidos tienen referentes comunes en un lenguaje de contención sobre la evaluación educativa más allá de lo instituido por el entorno inmediato del INEE.

Este lenguaje se modela a través de un conjunto de discursos⁵, de objetos y de acciones, cuyos contenidos y formas constituyen un sistema de convenciones compartidas que posibilitan las interconexiones, identificaciones, distinciones, la asignación de atributos de las situaciones en las que se participa.

A continuación expongo algunos aspectos relevantes de los discursos, objetos y acciones del lenguaje de contención de los comités del INEE, que permiten ampliar la comprensión sobre la evaluación educativa como fenómeno socio cultural.

3. EL DISCURSO SOBRE LA EVALUACIÓN EDUCATIVA

El lenguaje de contención de la evaluación educativa está fuertemente influenciado por los discursos oficiales de las autoridades en el campo educativo que cuentan con el respaldo del marco jurídico mexicano que tematiza la evaluación educativa. El Cuadro 3 resume las leyes, los acuerdos, los programas y los derechos principales que constituyen este marco jurídico. Disponer de una normativa específica y de un presupuesto asignado para llevar las acciones consecuentes, son indicios del nivel de institucionalización que ha adquirido el tema de la evaluación educativa y su importancia relativa en la agenda de políticas públicas.

El referente aprobado para nombrar y actuar en el terreno de la evaluación educativa son un conjunto de enunciados que se interrelacionan en un sistema conceptual y axiológico. Este referente junto con otras convenciones implícitas en el comportamiento cotidiano al realizar evaluaciones educativas se ligan⁶ configurando un esquema de percepción sobre la calidad educativa en general y la evaluación educativa en particular, como un medio para verificar el cumplimiento del derecho público a la educación y proporcionar información valiosa para superar las fallas que impiden que las personas ejerzan este derecho elemental.

⁵ Me refiero a las prácticas comunicativas en el campo educativo, tanto lingüísticas como no lingüísticas.

⁶ Dice María Guadalupe Mora Pizano: "La globalización, es el contexto del conjunto de políticas gubernamentales que se han desarrollado desde entonces, y por tanto, ... las políticas educativas tienen su origen en las condiciones derivadas del reordenamiento mundial de las economías, del surgimiento de bloques económicos regionales y la mundialización de los mercados que han colocado como ideales de logro la "calidad", la "productividad", la "rentabilidad" y la "competitividad", en el Tercer Congreso Nacional y Segundo Internacional "Retos y Expectativas de la Universidad" realizado en el 2002 en la Universidad Autónoma del Estado de México, http://www.congresoretosyexpectativas.udg.mx/Congreso%203/Mesa%201/Mesa1_11.pdf consultado el 19-02-2009.

CUADRO 3. RESUMEN DEL MARCO JURÍDICO MEXICANO SOBRE EVALUACIÓN EDUCATIVA

El artículo 3º. de la Constitución.	La educación como un derecho. La educación que imparta el Estado tenderá a desarrollar armónicamente todas las facultades del ser humano y fomentará en él, a la vez, el amor a la Patria y la conciencia de la solidaridad internacional, en la independencia y en la justicia. La educación es laica, científica.
La ley General de Educación.	Art. 12, 29, 30 y 31 precisan: -Responsabilidad exclusiva de la Secretaría de Educación Pública de evaluar el sistema educativo nacional. -Lineamientos para guiar las evaluaciones locales. -Carácter sistemático y permanente para que resultados sean base de decisiones. Art. 50 define que "la evaluación de los educandos comprenderá la medición en lo individual de los conocimientos, las habilidades, las destrezas y, en general, del logro de los propósitos establecidos en los planes y programas de estudio". Indica que las instituciones informarán periódicamente a educandos y padres de familia los resultados y calificaciones de exámenes y observaciones sobre el desempeño que permita mejorar su aprovechamiento. Art. 69, 70, 71 asigna a Consejos de Participación Social la función de conocer los resultados de las evaluaciones.
Acuerdo No 200 (1994)	Normas para la evaluación del aprendizaje en educación primaria, secundaria y normal. Considera la evaluación permanente y sistemática para la adecuación de los procedimientos educativos, la promoción de los educandos, el diseño y actualización de planes y programas de estudio y la planeación del sistema educativo nacional. Art. 1. Obligatoria Art. 2. A lo largo del proceso con procedimientos pedagógicos adecuados. Art. 3 Llevar a decisiones oportunas para la eficiencia de la enseñanza y del aprendizaje Art. 4. Congruencia entre evaluaciones y propósitos del programa. Art. 5 al 11 Especificaciones de las calificaciones. Otras disposiciones: promoción de grado, acreditación de estudios y regularización.
Acuerdo 286	Lineamientos para la revalidación de estudios hechos en otros países y para la acreditación de conocimientos adquiridos en forma autodidáctica o por la experiencia laboral.
Acuerdos No 243 y 279	Bases para el otorgamiento de reconocimiento de validez oficial de estudios institucionales particulares.
Ley Federal de Acceso a la Información Pública Gubernamental del 2002	Promueve apertura de información, manteniendo confidencialidad sobre información relativa a personas individuales.
La Ley de Información Estadística y Geografía (1980)	Considera las escuelas como unidades de información.
Programa Nacional de Educación 2001-2006	La educación ha tenido un crecimiento explosivo, en un contexto de alta desigualdad y grandes zonas de pobreza. Hay fallas técnicas, administrativas y políticas. Grandes objetivos: la calidad del proceso y logro, equidad y justicia social y rendición de cuentas- reforma de la gestión institucional, además de la cobertura. Enfoque educativo para el siglo XXI (hasta el 2025) por niveles educativos. Federalización, financiamiento, mecanismos de coordinación, consulta y participación, evaluación, investigación e innovación educativa. Calidad apreciada por la evaluación con mecanismos rigurosos y confiables, independiente de las autoridades, cuyos resultados serán difundidos y utilizará para el mejoramiento. Satisfacer criterios básicos de calidad de todo tipo de instituciones. Operar un sistema de información moderno y eficiente para conocer situación de aprendizaje, niveles de desempeño de maestros, escuelas, entidades. Sistema nacional de indicadores educativos para la planeación. Para consolidar el sistema nacional de evaluación educativa se crea el INEE cuyas funciones se articulan con otras instancias responsables en la SEP y entidades federativas.
Decreto de creación del INEE (2002)	Artículo 5 fracción I: definir instrumentos de manera conjunta con Secretaría de Educación Pública, una política nacional de evaluación que contribuya a la elevación de la calidad de la educación. Autonomía de autoridades, pero con una relación permanente y estrecha (cercanía) para asegurar que los resultados de la evaluación se utilicen para sustentar las decisiones. Orientaciones: alcanzar la mayor calidad en el desarrollo de modelos e instrumentos de evaluación (confiables y válidos).

FUENTE: Cuadro elaborado por la autora con base en la información INEE (2004) "2.Marco Normativo, principios y orientaciones" en *Documentos Institucionales del Plan Maestro de Desarrollo 2004-2012*, pp. 112-124

De manera directa la Ley General de Educación Art. 50 define la unidad de análisis de la evaluación educativa: "la medición en lo individual de los conocimientos, las habilidades, las destrezas y, en general, del logro de los propósitos establecidos en los planes y programas de estudio". Y señala la obligación de las instituciones de informar periódicamente a educandos y padres de familia los resultados y calificaciones de exámenes y observaciones sobre el desempeño que permita mejorar su aprovechamiento.

A partir del análisis de discurso con enfoque semiótico del marco normativo se elaboró el Mapa conceptual 1 que registra el sentido práctico⁷ de la evaluación en las fuentes normativas, que la establecen como estrategia para contrarrestar las condiciones que han provocado "fallas". En los discursos la situación actual de la educación es una distorsión o "anormalidad", de donde se infiere que hay un presupuesto: existe una condición adecuada o normal del sistema educativo nacional que es la Calidad educativa.

La calidad educativa es entendida en términos de gestión institucional, donde la eficiencia del proceso y de los resultados educativos es apreciada por la evaluación con mecanismos rigurosos y confiables, independiente de las autoridades, y la información obtenida será difundida y utilizada para el mejoramiento del sistema educativo nacional.

MAPA 1. MARCO CONCEPTUAL DE LA EVALUACIÓN EDUCATIVA



FUENTE: elaborado por la autora.

⁷ Utilizo la noción de "sentido práctico" tal y como la propone Pierre Bourdieu: se trata de un hábito incorporado que proporciona a los agentes sociales una disposición hacia la acción frente a las situaciones cotidianas a las que se enfrentan, sin tener que estar reflexionando y calculando todos los pormenores de la situación, sino utilizando los recursos de este aprendizaje social (Bourdieu, 1972, 1980, 2002).

Por un efecto de espejo, la evaluación educativa en el discurso oficial es también una manera de cuestionar a las administraciones públicas pasadas y a las condiciones asociadas con una cultura de irresponsabilidad e impunidad⁸. Como parte de las propuestas de solución se establece que la responsabilidad de la calidad educativa no es monopolio del gobierno, sino que supone la intervención de otros agentes sociales, entre los cuales juega un importante papel las instituciones académicas como el agente social investido de los atributos de rigor y confiabilidad científica que dan legitimidad a la información obtenida a través de indicadores de carácter universal, promovidos por las instituciones financieras internacionales tales como el Banco Mundial y la Organización para la Cooperación y el Desarrollo Económico, con los cuales se mide la eficiencia del sistema educativo.

El discurso oficial sobre la evaluación educativa, deja de circunscribirse a la evaluación tradicional de los alumnos como una práctica individual de los maestros orientada a la aprobación o no de los alumnos, para ser sustituida o complementada por otras formas de hacer en donde la aplicación de pruebas estandarizadas es una característica dominante. Una cultura de evaluación educativa con este enfoque coloca la evaluación como un aspecto clave del principio de rendición de cuentas y participación social que guía a la política pública de desarrollo social y humano.

Las responsabilidades del INEE que aparecen resumidas en el Cuadro 4, exhiben el carácter operativo y normativo conferido al Instituto. Al relacionar los enunciados del discurso del marco jurídico sobre la evaluación educativa se aprecia que este contenido se describe como una cuestión técnica, válida y científica. El análisis del sentido de este supuesto permite observar una cadena de ideas asociadas que parte de una matriz dicotómica: lo científico y lo que no lo es, acreditando a unos agentes portadores de lo primero, entre quienes está el INEE y sus comités de expertos, y desacreditando a otros agentes portadores de lo segundo, quienes elaboran y aplican exámenes cuyo proceso y resultados no están documentados ni son difundidos.

CUADRO 4. RESPONSABILIDADES DEL INEE

DISEÑAR	DESARROLLAR	DIFUNDIR	VINCULAR
Sistema de indicadores de la calidad del Sistema Educativo Nacional	Evaluaciones del aprendizaje	Resultados	Colaborar con la Secretaría de Educación Pública (SEP) en la evaluación del Sistema Educativo Nacional
Instrumentos y sistemas de evaluación	Modelos para la evaluación de las escuelas	Cultura de la evaluación	Representar a México ante los organismos internacionales de evaluación educativa
	Extensión de la evaluación en las entidades y la evaluación de programas y proyectos prioritarios		Coordinar la participación del país en los proyectos internacionales
	Acciones de capacitación		
	Investigaciones en la materia		

FUENTE: Elaborado por la autora con base en los documentos publicados en la página WEB del INEE www.inee.edu.mx

El discurso de la evaluación educativa sigue un principio de disyunciones, por ejemplo desigualdad contra equidad, conveniente contra no conveniente, pero no aclara que la situación que guarda actualmente la evaluación educativa es resultado de condiciones sociales e históricas donde se viven conflictos y

⁸ Lo cual ha generado un campo de tensiones entre la Secretaría de Educación Pública y el Sindicato Nacional de Trabajadores de la Educación, y al interior de éste.

acuerdos en un proceso con una dinámica cambiante a través de la cual se hace institucionalizable, organizable, regulable, administrable un esquema de percepción donde el referente “científico” no se utiliza sólo con el propósito de ampliar el conocimiento sino que también representa un arma para imponer una actitud autoritaria cimentada en “la verdad del conocimiento científico”, lo cual es opuesto al espíritu científico que reconoce que todas las afirmaciones, incluso las basadas en criterios experimentales, conllevan una dosis de incertidumbre.

El proceso y resultados de los comités de evaluación del INEE representan una oportunidad de avanzar hacia una comprensión razonable del aprendizaje y su contexto, con la posibilidad de esquivar imperativos dogmáticos en la medida que los procedimientos técnicos sean documentados y problematizados, no estereotipados como fuentes de conocimiento sino como estrategias para detectar y eliminar errores, a sabiendas que la precisión absoluta es inasequible⁹ y carece de propósito si no aporta valor para responder a la problemática educativa.

4. LOS OBJETOS DE LA EVALUACIÓN CUALITATIVA

Se ha descrito la dinámica del trabajo colegiado de los comités para que el INEE cumpla una de sus funciones principales: el diseñar pruebas nacionales de aprendizaje. El resultado del proceso, es decir los Exámenes de la Calidad y el Logro Educativos (Excale) representan en sí mismos una influencia importante para modelar el sentido social y cultural que tiene la evaluación educativa en México actualmente. Estos exámenes son instrumentos para conocer el tipo de habilidades y conocimientos que dominan los estudiantes, su inserción en el currículum nacional, la importancia cognoscitiva para el aprendizaje de algunas asignaturas que se consideran fundamentales; permiten establecer una tipología de preguntas y la dificultad que tiene el contenido para los distintos grupos de estudiantes del Sistema Educativo Nacional.

En el Cuadro 5. *Cobertura de los Excale*, se precisa el alcance de este objeto de la evaluación educativa, los cuales forman parte de la agenda de la política educativa, demandan recursos altos para disponer del trabajo y de información de especializada, de manera sistemática a través del tiempo; para poder dar una respuesta muy sensible a cambios curriculares que imponen un realineamiento de examen mediante una acción permanente de los comités que analizan el currículo nuevo, elaboran nuevas especificaciones, crean nuevos reactivos y pruebas empíricas que validen los modelos de examen adecuados.

Estos objetos de evaluación educativa se presentan como la concreción de metodologías con un alto grado de complejidad técnica basada en el paradigma de la ciencia experimental y en el grupo de expertos que la diseñan y desarrollan, entre los cuales se cuentan los miembros de los comités. Bajo estas condiciones los exámenes se asumen como productos profesionales, neutrales y deseables. La racionalidad del proceso de diseño, desarrollo y aplicación de estos objetos proporciona confiabilidad y validez a la información generada.

⁹ Karl Popper habla del método de “falsación” como parte de la lógica del conocimiento científico que no puede comprobar los resultados obtenidos en el proceso de investigación como verdad absoluta, sino someterlos a prueba, para saber si se sostienen de acuerdo a los condicionantes explicitados, en sus obras *La lógica de las ciencias sociales* editada por Grijalbo y *En búsqueda de un mundo mejor* de la Editorial Paidós.

CUADRO 5. COBERTURA DE LOS EXCALE

Tipo	Alumnos	Período cubierto ¹⁰	Áreas evaluadas	Complemento
EXCALE-00	3° Preescolar	Al finalizar el ciclo escolar 2006-2007	- Lenguaje y comunicación - Pensamiento matemático	Respuestas de alumnos, padres de familia, docentes y directores sobre aspectos del contexto en donde se desenvuelve la labor educativa.
EXCALE-03	3° Primaria	Al finalizar el ciclo escolar 2005-2006	- Español - Matemáticas - Ciencias Naturales - Ciencias Sociales	
EXCALE-06	6° Primaria	Al finalizar el ciclo escolar 2004-2005	- Español - Matemáticas	
EXCALE-09	3° secundaria	Al finalizar el ciclo escolar 2004-2005		

FUENTE: Resumen de la información publicada en la página WEB del INEE www.inee.edu.mx consultada en 15 febrero 2009, elaborado por la autora.

El enfoque de esta racionalidad considera el uso de la estadística para el análisis e interpretación de los resultados de los Excale como el referente principal para construir la imagen de que los niveles de competencia alcanzados miden racional y objetivamente, con una base empírica y según una lógica deductiva, la problemática educativa. En el Cuadro 5 se muestra el alcance de los Excale.

El método utilizado se apega a categorías que apoyan su carácter universal en las rutinas referidas en la situación de comunicación de los comités, subrayando el seguimiento de lineamientos y estándares precisos de acuerdo a diferentes asociaciones especializadas: American Educational Research Association (AERA), American Psychological Association (APA), National Council on Measurement in Education (NCME)¹¹, Standards for Educational and Psychological Testing de AERA; Estándares de Calidad para Instrumentos de Evaluación Educativa del Centro Nacional de Evaluación de México CENEVAL¹², Referenced National Examinations for Certification and Selection of Students de la Asociación para el Estudio de la Evaluación Educativa en Sudáfrica (ASSESA). Lo mismo pasa con el modelo de construcción de pruebas, de tipo criterial alineado al currículum, que a nivel internacional y nacional es dominante.

Frente al carácter universal de estos objetos de la evaluación educativa de gran escala, en los comités se subraya el carácter complejo del objeto de estudio de los Excale: el aprendizaje. El aprendizaje es un proceso que debe contextualizarse para tener una comprensión clara. Los resultados obtenidos y difundidos son las medias, los índices, coeficientes, que son abstracciones que responden a las categorías e indicadores determinados con base en los lineamientos y estándares mencionados. El tratamiento cuantitativo de la información cumple con un principio de economía de la investigación, al priorizar resultados generalizables a todo el universo de estudio, estos resultados se integran a la contabilidad

¹⁰ La aplicación del instrumento es periódica cada tres años con el propósito de que los resultados obtenidos sean comparables y pueda hacerse un seguimiento generacional.

¹¹ NCME es una organización de la sociedad civil de Estados Unidos cuyo propósito es fomentar la mejora de la educación a través de avanzar la ciencia de la medición en el campo educativo, mejorar los instrumentos y procedimientos para la administración, contabilización e interpretación de la evaluación de programas educativos. Se orienta hacia la difusión de teorías, métodos, técnicas, procedimientos disponibles para valorar las características humanas, institucionales y sociales relevantes para el logro educativo individual y grupal.

¹² CENEVAL es una asociación civil sin fines de lucro cuya actividad principal es el diseño y aplicación de instrumentos de evaluación de conocimientos, habilidades y competencias, así como el análisis y la difusión de los resultados que arrojan las pruebas siguiendo procesos estandarizados de diseño y construcción y se apegan a las normas internacionales.

Desde 1994 se constituye en la Ciudad de México por instituciones educativas, asociaciones y colegios de profesionales, organizaciones sociales y productivas y autoridades educativas gubernamentales. Sus actividades se sustentan en los últimos avances e investigaciones de la psicometría y otras disciplinas, así como en la experiencia y compromiso de su equipo, integrado por casi 400 especialistas provenientes de las instituciones educativas más representativas del país y organizaciones de profesionales con reconocimiento nacional.

social para medir y ubicar aciertos y errores de las políticas educativas y apreciar sus alcances: por abajo del básico, básico, medio, avanzado, como datos exteriores sin sesgos de carácter partidistas, personales, o de cualquier otra índole.

El criterio técnico de eficacia de los Excale es el apego al currículo establecido en el sistema educativo nacional, pero este criterio influye para que los objetos de la evaluación educativa mantengan los referentes tradicionales enciclopedistas consignados en el currículo, que poco tienen que ver con la trayectoria de las niñas y los niños y jóvenes estudiantes como lectores y escritores. Las obras leídas, las opiniones personales, las conversaciones, el acceso al contexto de la producción, la construcción de sentido de cada persona se aleja mucho del currículo determinado por las instituciones educativas. En el Cuadro 6 se pueden apreciar las diferencias en la caracterización de las prácticas de escritura de las niñas, los niños y jóvenes (prácticas vernáculas), y la caracterización de las prácticas dominantes en el currículo escolar.

CUADRO 6. PRÁCTICAS DE ESCRITURA DE NIÑAS, NIÑOS Y JÓVENES EN LA ERA DIGITAL

	Prácticas extraescolares	Prácticas escolares dominantes
Dinámica	Cambiante, efímera, flexible, versátil	Rígida, "seria", acotada
Difusión	Libre en redes personales, alta frecuencia	Impuesta en un sistema institucional de acceso general, de frecuencia estandarizada
Planeación	Autogeneradas, abierta, intencionalidad autoregulada	Preestablecidas, impuestas y reguladas externamente
Cobertura	Personal, familiar, amistad	Escolar
Antecedente	Aprendidas informalmente	Aprendidas formalmente como parte del currículo
Ejemplos	Conversaciones, mensajes, diario íntimo, diario de viaje, postales, cartas, novelas de moda para la niñez y adolescencia, revistas, contabilidad, fotos, canciones, foros, chats, radioescucha, teleaudiencia, videojuegos, interpretación de música, listas de compra, espectaculares en la vía pública, servicios bancarios, entre otros	Apuntes de una asignatura, tareas extraclase, comentario de texto, exámenes, redacción.
Tecnologías de información y computación	Teléfono celular, webs abiertas, computadoras personales, lápiz, pluma, papel	Webs institucionales, computadoras institucionales

FUENTE: Elaboración propia a partir de los resultados de la etnografía y sondeo de percepción de estudiantes de 1ero., 3º y 6º de la primaria pública Escuela Catorce de Abril de la ciudad de México realizado en 2007.

Más allá de la contabilidad social, los Excale relevan información valiosa para mirar más profundamente aspectos clave del proceso de aprendizaje, aspectos ligados con las competencias para la vida y con la construcción de la ciudadanía, alineados con el enfoque de las políticas públicas. Este tipo de información complementaria a la estandarizada, se acumula, pero su organización, mantenimiento, clasificación y análisis es costosa, requiere de una significativa inversión de tiempo del personal del INEE que debe cumplir con otras responsabilidades, y el procesamiento de información exige competencias técnicas especializadas y considerar no sólo las prioridades derivadas de entender a la educación como un fenómeno de gestión pública, sino como un fenómeno de construcción de identidades relacionadas con las prácticas lectoras y escritoras que han variado en la era digital, porque cambian los espacios en los que tienen lugar, los soportes materiales, los contenidos y la manera de abordarlos para leerlos y/o producirlos por parte de las niñas, los niños y jóvenes de asisten a la escuela pero que aprenden significativamente en el entorno extraescolar.

Esta situación obstaculiza el aprovechamiento práctico de las lecciones aprendidas de la operación de la evaluación y limita la efectividad del principio de participación social y rendición de cuentas, por la falta de una gestión del conocimiento. El costo por desinformación puede llevar a: una comprensión limitada al currículo de las condiciones y dinámica de la calidad educativa, una imposibilidad de ampliar la masa crítica de conocimientos para tomar decisiones fundamentadas y mejorar la política educativa.

Las niñas, los niños y jóvenes que estudian el nivel básico del sistema educativo nacional, sí leen y escriben, pero lo hacen de una manera no apegada al currículo escolar. Sus prácticas son de carácter informal dentro de las comunidades de práctica a las que pertenecen: chats, foros, fotologs, mensajes telefónicos, entre otros, cuyo propósito es expresarse, divertirse, conocerse mejor, sin un interés académico, integran su capacidad de ser autores de escritos, fotografías, ortotipografías¹³ y de lecturas que se integran a la dinámica vivencial de estar acompañados en diferentes tipos de comunidades.

El análisis de las “evidencias” de las respuestas de los Excale, bajo un tratamiento cualitativo de casos permitiría rescatar elementos de las prácticas de las niñas, los niños y jóvenes, siguiendo la línea de investigación de Daniel Cassany¹⁴, para identificar: el nivel de complejidad; el nivel de consistencia y coherencia que ayudara a problematizar la idea de que a los niveles bajos de rendimiento corresponden respuestas caóticas o simples; caracterizar el sentido de la identidad estudiantil. Estos objetos de evaluación educativa pueden proporcionar indicios sobre las formas en que se interrelacionan las prácticas propias de las niñas, los niños y jóvenes, con las dominantes en la escuela, lo cual representa una oportunidad de encontrar opciones para fortalecer el aprendizaje significativo de la lectura y escritura, como parte fundamental de las competencias ciudadanas.

Los objetos de evaluación educativa se conforman al currículo y paradójicamente guardan elementos de aspectos ignorados políticamente porque no se asigna presupuesto para sistematizarlos y difundirlos, que tienen que ver con las prácticas no escolares de las niñas, los niños y jóvenes como espacios de aprendizaje. Los Excale son objetos utilizados hasta ahora para determinar el nivel de logro en función de salir bien en un examen, pero también abonan elementos para fortalecer el proyecto educativo más allá del ajuste a los estándares nacionales e internacionales establecidos sino implican una estrategia para mejorar sustancialmente el proceso de aprendizaje con propuestas innovadoras que rescaten elementos de prácticas cotidianas y significativas para mediar el currículo. La construcción de los Excale está en un camino donde la integración de las tecnologías de información y computación, el desarrollo de multitareas multimodales significativas para las y los estudiantes en el contexto de la era digital, serán parte de su reto.

5. LAS ACCIONES DE EVALUACIÓN EDUCATIVA

Una de las acciones claves para asegurar el funcionamiento efectivo y eficiente del proceso de diseño, construcción y validación de los Excale es la capacitación dirigida a los comités de especialistas y docentes que participan en el proceso. Se trata de una estrategia de profesionalización para calibrar competencias técnicas y una estrategia de participación en el proyecto de evaluación. La capacitación se

¹³ Usos y convenciones particulares por medio de elementos tipográficos con símbolos, colores, fuera del espacio de diccionario tradicional.

¹⁴ Profesor de Análisis del Discurso en la Universidad Pompeu Fabra (Barcelona), licenciado en Filología Catalana y doctor en Filosofía y Letras que ha publicado sobre comunicación escrita y enseñanza de la lengua y en México ha colaborado en programas de promoción de la lectura y la escritura.

proporciona con las conferencias que inauguran los trabajos de los comités y los talleres para refinar las competencias técnicas con base en la adopción de principios formalizados y estandarizados a nivel internacional.

Aunque la duración de la capacitación es corta, la capacidad de fijación de los contenidos y formas de evaluar se potencia en la medida que la información transmitida está preservada en manuales que pueden ser usados posteriormente.

Durante la capacitación la forma y contenido de la evaluación se presenta como algo dado desde la autoridad que confiere el prestigio internacional del INEE al realizar pruebas de aprendizaje a gran escala bajo parámetros internacionalmente convenidos. La capacitación recibida se convierte en un signo de distinción en el curriculum vitae de los participantes que los apoya como estrategia de promoción del profesionalismo y especialización.

El flujo de información que sigue el proceso de capacitación combina la comunicación unidireccional desde los representantes del Instituto hacia los miembros de los comités, con una interconexión entre los mismos derivada de la puesta en común de problemáticas compartidas en las experiencias personales en el campo educativo, que propiciará la formación de comunidades con capacidad de influencia en el interés público.

El manual de los comités proporciona las herramientas necesarias para llevar a cabo la tarea comprometida, con un enfoque técnico altamente estructurado que delimita el qué y por qué, el cómo, el cuándo, el cuánto, el quién. El manual integra en un único mecanismo los elementos constitutivos del tema de manera concisa, con una redacción precisa de acuerdo a regulaciones y tendencias convencionales a nivel internacional y con una presentación que garantiza el rápido acceso a la información específica, a través de subáreas temáticas bien definidas y la comprensión sistemática del público destinatario mediante la utilización de esquemas como diagramas de flujo, árboles de decisión, entre otros. La fundamentación de su pertinencia es la forma participativa y consensuada, dentro del sector de especialistas sobre las propiedades claves de las buenas prácticas evaluativas.

El manual incluye la definición de conceptos y procedimientos, los pasos clave en la aplicación de métodos. Contiene un conjunto de anexos para ampliar el conocimiento, los términos de referencia y ejercicios de aplicación. El manual se caracteriza por una alta fijación en la memoria por su durabilidad y por la capacidad de guardar información para preservarla y usarla posteriormente; y una participación selectiva de los individuos que emplean el medio. Para usar el manual se aplican diferentes habilidades, facultades y recursos para codificar y decodificar los contenidos y su forma.

Otra acción relevante de los comités es el proceso de argumentación a través del cual realizan la tarea encomendada, la cual es su razón de ser. La existencia de los comités es efímera, su nombramiento dura el tiempo que tardan en elaborar los productos esperados. Para cumplir su objetivo los miembros del comité tienen que seguir unas reglas establecidas por el INEE y otras construidas en la situación de comunicación específica. Reconocer, seguir y valorar las reglas de argumentación implica un descentramiento de la posición que cada uno de sus integrantes sostiene, un escuchar con atención la exposición de información ampliada, más allá de lo que se transmite a través de los medios técnicos institucionalizados.

En la medida que son tan diversos los tipos de agentes en contacto, la cultura de evaluación no sigue un patrón lineal repetitivo impuesto unidireccionalmente, sino que la cooperación prevalece en la medida

que haya una articulación con los intereses de los integrantes y las estructuras a las que pertenecen, que van más allá de la cultura de evaluación.

A través de las prácticas del comité se define su identidad, permeada por el lenguaje de contención que conecta con el papel prescrito de los comités en el marco normativo, por los atributos que los otros miembros del comité otorgan a cada uno de sus integrantes en la situación de comunicación, y por la autopercepción. Un rasgo común de la identidad de los miembros del comité es que tienen la responsabilidad de:

- superar sus desacuerdos y de llegar a puntos de convergencia;
- sostener lo que acuerdan ya que los productos entregados por un comité son firmados por todos y cada uno de sus integrantes;
- satisfacer la mayoría de sus intereses;
- quedar en condiciones de repetir la experiencia y ser invitados a formar parte de otro comité en futuras oportunidades.

Estas prácticas retan a los integrantes a mantener el interés, controlar la inquietud de exteriorizar lo que piensan, saben o han vivido, frente a la regla de respetar turnos para tomar la palabra, disentir sin que deje de apreciarse el valor de las necesidades e intereses del otro; aclarar los paradigmas e incluso prejuicios personales que pueden obstaculizar el llegar a un enfoque colectivo. Este asunto es parte de asegurar la confiabilidad y validez de que los productos elaborados son consistentes y reproducibles.

Como resultado se experimentan diferentes niveles de cambio en la autopercepción y en la consideración institucional, de lo cual de manera hipotética se puede inferir que el aprendizaje generado por las prácticas del comité, las competencias, podrían ser transferibles hacia otros contextos.

6. LECCIONES APRENDIDAS

La evaluación del aprendizaje escolar es un fenómeno complejo que constituye un tejido de interacciones, determinaciones, azares, difíciles de comprender sin hacer asociaciones o separaciones a priori, a menos que pensemos bajo los tres principios epistemológicos sugeridos por Edgar Morin: el principio dialógico que nos recuerda que la dualidad mantiene a la unidad; el principio de recursividad que nos invita a reconocer que cada elemento de un fenómeno es producido y al mismo tiempo, productor; y el principio hologramático que ayuda a mediar la idea de que cada una de las partes del fenómeno están en el todo y que el todo está en cada parte.

A manera de síntesis y conclusiones de la reflexión en torno a las prácticas de los comités del INEE como fenómeno sociocultural de participación social y rendición de cuentas en el campo educativo se enuncian y describen las siguientes paradojas que pretenden distinguir y conjuntar las características descritas, en una aproximación no determinista, sino atenta y responsable.

Paradoja 1

La estructura de los comités de evaluación del estudio de caso refleja por un lado, una regulación estricta de su actuación, es decir es algo ya instituido; y por otro lado la intervención de todos y cada uno de los

miembros de los diferentes comités imprime una dinámica específica que rebasa lo instituido con elementos emergentes – instituyentes.

Paradoja 2

La situación de comunicación de los comités presenta una paradoja entre la ilusión¹⁵ de que lo instituido es lo natural, la “ilusión de transparencia” como un “debe ser así”, para que la evaluación sea un ejercicio técnico de calidad, y la actualización de valores y símbolos que vinculan a los participantes más allá de las premisas técnicas hacia intereses públicos que se intensifican, que no son naturales sino que están social y históricamente determinados.

Paradoja 3

La paradoja del discurso sobre la evaluación educativa es que se articula con la objetivación del proceso de trabajo de los comités, con los documentos producidos y puestos en circulación, la opción de establecer distancia sobre lo que ocurre en la situación de comunicación particular, y ampliar la visión sobre el fenómeno evaluativo para comprender los mecanismos para producir los resultados de evaluación y cómo aprovecharlos en otros espacios y tiempos. Estos dos componentes se estructuran en una suerte de espiral de crecimiento donde la evaluación abre la posibilidad de ser evaluada ella misma (meta evaluación más allá de los procedimental), y con ello ser problematizada y generar aprendizajes significativos para los miembros del comité de evaluación y para otros agentes sociales involucrados, que van más allá de sólo repetir lo transmitido.

Paradoja 4

Los objetos de los comités en su contenido y forma, contribuyen a dar credibilidad al esquema de percepción dominante sobre la evaluación educativa. Sin embargo en situaciones de tensión entre grupos de interés, tales como la asignación de presupuestos a las escuelas con base en los resultados de la evaluación del aprendizaje, o ligar los estímulos al desempeño docente, entre otros mecanismos, se pone de manifiesto una dimensión que el discurso y los objetos de evaluación educativa soslayan: el campo de fuerzas donde los involucrados en el campo educativo ocupan posiciones distintas que representan un acceso diferencial a recursos económicos, políticos y de reconocimiento social.

Paradoja 5

La dinámica de las acciones del comité incluye cooperación para lograr entregar un resultado consensuado, y también cierto grado de competencia al convivir perspectivas individuales diferentes ante una tarea interdependiente.

Las paradojas de los comités de evaluación educativa como fenómeno sociocultural donde se encuentran representadas muchas voces, enlazan tanto mecanismos de control como puntos de ruptura y oportunidades que fomentan la producción de significados plurales de la evaluación, sin dogmatismo totalizadores, como una cuestión técnica pero también como una cuestión sociocultural de participación social y rendición de cuentas, que permiten pensar en:

¹⁵ La “ilusión de la transparencia” (Bourdieu, 1975:29) es producto de la familiaridad de los agentes de este campo con las rutinas -sus contenidos y formas- en las que participan que funcionan como lo que William Roseberry llama “lenguaje de contención”, el cual opera como marco de referencia para vivir a través de y actuar hacia los ordenamientos sociales de la evaluación educativa (Roseberry, 1994).

- explorar, contrastar, intercambiar y alimentar propuestas de aproximaciones particulares hacia la mejora de la educación, que hacen de las acciones y los datos instrumentos tangibles que posibilitan la comprensión de la realidad, sin caer en dogmatismos totalizadores ni simplistas;
- abrir opciones de gestión educativa como un asunto de interés público, no como un hecho dado, sino como un proceso que cambia y se intensifica, a lo largo de un proceso de negociación de sus significados sociales;
- nutrir un horizonte ético para crear influencia y establecer relaciones de equidad en el campo educativo como una responsabilidad y exigencia moral de tener que responder para construir condiciones de vida buena.

REFERENCIAS BIBLIOGRÁFICAS

- Backhoff, E.; Monroy, L.; Peon, M.; Sánchez Moguel, A.; y Tanamachi M. L. (2005). *Exámenes de la Calidad y el Logro Educativos (I aprendizaje del Español y las Matemáticas en la educación básica en México. Sexto de primaria y tercero de secundaria*. México: INEE.
- Andrade, A. (2006). *Evaluación de impacto y política social I*. México Universidad Iberoamericana: Instituto de Desarrollo Social, Fundación ProEmpleo A.C.
- Andrade, A. (2008). Diálogo en el aula. Revista *Didac*, 51, Primavera 2008, pp.27-31.
- Bourdieu, P. (1975). *El Oficio de sociólogo*. México: Siglo XXI.
- Bourdieu (1972). Structures, habitus et pratiques. *Esquisse de'une theorie de la pratique*. pp.174-189. París: Droz.
- INEE (2008). *Manual técnico de especificaciones de reactivos*. México: INEE.
- INEE (2004). 2. Marco Normativo, principios y orientaciones. *Documentos Institucionales del Plan Maestro de Desarrollo 2004-2012*, <http://www.inee.edu.mx> consultado el 19-02-2009.
- INEE (2005). *Plan General de Evaluación del INEE 2005-2016*, <http://www.inee.edu.mx> consultado el 19-02-2009.
- Mora Pizano, M. G. (2002). Retos y Expectativas de la Universidad. *Tercer Congreso Nacional y Segundo Internacional*, México, Universidad Autónoma del Estado de México, http://www.congresoretosyexpectativas.udg.mx/Congreso%203/Mesa%201/Mesa1_11.pdf consultado el 19-02-2009.
- Morin, E. (2001). *Introducción al pensamiento complejo*. Madrid: Gedisa.
- Roseberry, W.(1994). Language of contention. En Joseph y Nugent (editors), *Everyday forms of state formation*, pp.355-366. Durham: Duke University Press.

ANÁLISIS MULTINIVEL DE LOS FACTORES ASOCIADOS A LA EDUCACIÓN CÍVICA EN ESTUDIANTES DE SECUNDARIA

Joaquín Caso, Eduardo Hernández y Juan Carlos Rodríguez

Revista Iberoamericana de Evaluación Educativa 2009 - Volumen 2, Número 1

<http://www.rinace.net/riee/numeros/vol2-num1/art12.pdf>

Fecha de recepción: 10 de marzo de 2009
Fecha de comunicación de dictamen: 5 de abril de 2009
Fecha de aceptación: 5 de abril de 2009

El reciente interés por conocer los resultados de los programas y prácticas educativas asociadas con la formación cívica y valoral de los estudiantes, ha derivado en la participación de diversos países en evaluaciones internacionales a gran escala entre las que destacan el World Values Survey (Inglehart y Baker, 2000), el Latinobarómetro (Lagos, 2005) y el Civic Education Study (Torney-Purta y Henry-Barber, 2004). A diferencia de la activa participación que México ha registrado en la última década en evaluaciones internacionales orientadas a conocer los niveles de logro en dominios curriculares tradicionales, son escasos los registros de experiencias evaluativas orientadas a explorar el componente afectivo asociado a dichos aprendizajes.

A este respecto el Civic Education Study (CIVED), estudio desarrollado por la International Association for the Evaluation of Educational Achievement (IEA), ha permitido conocer las actitudes cívicas de estudiantes entre 14 años y 17 años de edad registrando a la fecha más de 90 mil participantes de 28 países, incluyendo a Chile y Colombia.

El estudio surge en 1999 como una necesidad de analizar los cambios que tuvieron lugar en las escuelas y en los sistemas políticos a nivel global y del interés por conocer los impactos de los programas diseñados para inculcar principios democráticos en los estudiantes y por conocer sus actitudes y percepciones sobre los procesos democráticos que les rodeaban (Torney-Purta y Henry-Barber, 2004). La aplicación del CIVED ha proporcionado información sobre aquello que los jóvenes piensan de la democracia, sus expectativas, su funcionamiento y los valores implícitos que la sustentan. A diez años de este primer estudio internacional, la IEA decidió instrumentar una estrategia evaluativa de naturaleza similar denominada International Civic and Citizenship Education Study (ICCS), estudio en curso en el que participan 35 países, incluyendo México.

Por su parte, el movimiento de investigación de Eficacia Escolar (*School Effectiveness*), centrado en analizar por qué algunas escuelas son más eficaces que otras, ha permitido conocer el efecto de las variables contextuales asociadas al logro educativo, orientadas tanto por los hallazgos de la investigación empírica existente, como por aproximaciones que incluyen variables cuya incidencia aún no se ha comprobado claramente (Murillo, 2003). En México, al igual que en otros países, se han registrado recientemente un número considerable de estudios a gran escala, en los que se ha explorado principalmente el efecto de variables contextuales sobre el aprendizaje de las matemáticas y el español en estudiantes de educación básica (Backhoff, Bouzas, Hernández y García, 2007; Backhoff, Bouzas, Contreras, Hernández y García, 2007; Valenti, 2007; Backhoff, Andrade, Sánchez y Peón, 2008).

Estos esfuerzos contrastan de manera importante con los estudios que exploran el efecto de variables contextuales sobre otro tipo de dominios, particularmente aquéllos relacionados con contenidos curriculares de naturaleza cívica y ética. A este respecto, con excepción de las evaluaciones realizadas en México por el Instituto Nacional para la Evaluación de la Educación en 2005 con una muestra nacional de estudiantes de tercer año de educación primaria (Backhoff, Andrade, Sánchez y Peón, 2007), nuestro país no cuenta con registro de acciones evaluativas sistemáticas con estos alcances.

Los primeros estudios realizados a nivel internacional que exploraron las relaciones entre el logro educativo y estos factores llegaron a conclusiones en el sentido de que la influencia de la escuela parecía ser menos importante que la influencia de factores del entorno familiar y social (Coleman et al., 1966; Hanushek y Luque, 2003), siendo el nivel socioeconómico una variable determinante de los resultados educativos.

Por el interés que ha despertado en la investigación contemporánea el estudio de la influencia del nivel socioeconómico en el logro educativo, se han utilizado diversos criterios para definir conceptual y operacionalmente este constructo. Así, el nivel socioeconómico se ha estimado mediante el uso de indicadores tales como el ingreso familiar y la ocupación, estatus laboral y nivel educativo de ambos padres (Willms, 2001; 2003; Duru-Bellat, 2004). Otras aproximaciones han utilizado un espectro de indicadores más amplio, desagregándose en posesiones culturales (número de libros en casa, compra del periódico, instrumentos musicales, enciclopedia, etc.), posesiones materiales (automóvil, un velero, casa de campo, etc.) e ingreso familiar (Hansen, Rosén y Gustafsson, 2001). A este respecto, recientemente se ha sugerido la conformación de una variable compuesta denominada *capital cultural escolar* constituida a su vez por un conjunto de variables tales como el nivel educativo de los padres, las expectativas educativas personales y familiares, el número de libros en casa, el acceso a internet y la frecuencia de asistencia al cine (Backhoff *et al.*, 2007).

Así, reconociendo la necesidad de contribuir a la generación de conocimiento en torno a los factores asociados con el desarrollo de actitudes cívicas en estudiantes de secundaria, y teniendo como finalidad estimar los efectos tanto de variables personales como de naturaleza escolar, en lo general, y de los efectos del nivel socio-económico y el estrato escolar, en lo particular, se describe el estudio realizado y se presentan sus resultados.

1. LA INVESTIGACIÓN DESARROLLADA

1.1. El método

1.1.1. Participantes

Participaron 4,111 estudiantes de tercero de secundaria pertenecientes a 43 escuelas de los cinco municipios de Baja California, entidad fronteriza ubicada en el noroeste de México, que se distribuyeron, en cuanto a sus principales características de la siguiente forma: a) el 51.6% de los participantes fueron mujeres y el 48.4% hombres; b) con edades fluctuando entre los 13 y los 17 años con una edad promedio de 14.9 años; c) el 37.3% originarios del municipio de Tijuana, 25.1% de Ensenada, 21.7% de Mexicali, 10.7% de Rosarito, y 5.2% de Tecate; d) el 57.1% de los estudiantes inscritos en secundarias generales, 27.9% en secundarias técnicas, 10.4% en secundarias privadas y 4.6% en telesecundarias; y e) el 77.9% estudiantes inscritos en el turno matutino y 22.1% en el vespertino. La selección de las escuelas fue determinada mediante el método de muestreo por conglomerados en dos etapas, con probabilidades proporcionales al tamaño, teniendo como unidad última de selección la escuela y como unidad de observación al estudiante de tercero de secundaria.

1.1.2. Instrumentos

Civic Education Study (CIVED). El CIVED se conforma por 154 reactivos con opciones de respuesta tipo Likert agrupados en trece dimensiones: democracia, ciudadanía, gobierno, confianza en las instituciones, nuestro país, oportunidades 1, oportunidades 2, sistema político, escuela, aprendizaje en la escuela, acción política 1, acción política 2 y salón de clases. Evalúa, en general, aspectos actitudinales relacionados con el conocimiento de los principios fundamentales de la democracia, el concepto de

ciudadanía, las percepciones sobre las responsabilidades de los gobiernos, el nivel de confianza en las instituciones, las actitudes y percepciones con respecto a la nación, equidad de género, la oportunidad percibida para grupos minoritarios, los derechos políticos de la mujer y las expectativas de participación futura en actividades políticas (Torney-Purta y Henry-Barber, 2004). Los reactivos están agrupados en las siguientes escalas:

TABLA 1. ESTRUCTURA DE LA ADAPTACIÓN DEL CIVED PARA BAJA CALIFORNIA

Dimensión	Descripción
A. La Democracia	Explora las opiniones de los jóvenes sobre el concepto de democracia (23 reactivos).
B. Ciudadanía	Explora qué tipo de conceptos asocia un estudiante a la configuración de un buen ciudadano (15 reactivos).
C. El Gobierno	Explora las percepciones de los estudiantes sobre las responsabilidades sociales y económicas del gobierno (12 reactivos).
D. Confianza en las instituciones	Explora el nivel de confianza que los estudiantes manifiestan hacia las principales instituciones públicas (14 reactivos).
E. Nuestro país	Explora las actitudes y sentimientos de patriotismo y nacionalismo en los estudiantes (13 reactivos).
F. Oportunidades 1	Explora la percepción de los estudiantes en torno a las oportunidades que existen en este país para los grupos vulnerables (6 reactivos).
G. Oportunidades 2	Explora la percepción de los estudiantes con respecto a las oportunidades y derechos que deberían tener ciertos grupos vulnerables (14 reactivos).
H. El sistema político	Explora las opiniones e intereses que tienen los estudiantes sobre diversos asuntos relacionados con el sistema político nacional (7 reactivos).
I. La escuela	Explora en qué medida los estudiantes participan en la dinámica escolar (7 reactivos).
J. El aprendizaje en la escuela	Explora las opiniones de los estudiantes en cuanto al tipo de aprendizajes que se promueven en la escuela (7 reactivos).
K. Acción política 1	Explora el interés actual de los estudiantes en temas políticos y el grado de exposición a medios masivos de comunicación (10 reactivos).
L. Acción política 2	Explora las expectativas de participación futura en acciones políticas (12 reactivos).
M. Salón de clases	Explora la percepción de los estudiantes del ambiente del aula y de las prácticas pedagógicas que utilizan sus profesores (12 reactivos).

Entre las propiedades psicométricas documentadas por la aplicación de este instrumento, destacan índices de bondad de ajuste interno y externo con valores de entre .50 a 1.50 en los 154 reactivos que lo conforman, la confirmación de la unidimensionalidad de sus trece escalas, índices de consistencia interna en el rango de .64 a .82, e índices aceptables de bondad de ajuste para la totalidad de escalas que se sometieron al modelamiento de ecuaciones estructurales (Caso, González-Montesinos, Rodríguez y Urias, 2008). Dado el número heterogéneo de reactivos que conforman las escalas, se requirió de su transformación en una métrica uniforme con una media de 10 y una desviación estándar con valor de 2, a fin de hacerles comparables entre sí.

Cuestionario de variables de contexto. Cuestionario conformado por 23 variables contextuales que exploran información de naturaleza económica, social y cultural vinculada con actitudes cívicas de los estudiantes (Rodríguez, Caso, Contreras & Urias, 2007). Las variables exploradas por la presente investigación incluyen el tipo de secundaria y una variable compuesta denominada nivel socioeconómico, construida al integrar las variables nivel educativo de los padres, expectativas educativas del alumno, número de libros en casa, posesión de bienes materiales (como reproductor de discos compactos, de DVD, computadora, consola de videojuegos, horno de microondas y automóvil), y acceso a servicios (línea telefónica y conexión a Internet).

1.1.3. Procedimiento

La aplicación de instrumentos requirió de las siguientes acciones: a) traducción y adaptación del CIVED Internacional; b) aplicación piloto del instrumento a estudiantes de tercer año de secundaria que confirmó la pertinencia y claridad en las instrucciones y reactivos; d) capacitación a 41 coordinadores de la aplicación, e) aplicación del instrumento a gran escala, f) supervisión de la aplicación, g) lectura electrónica de la hoja de respuestas, y h) preparación de la base de datos.

1.1.4. Análisis de datos

A fin de estimar el efecto de las variables de contexto en las actitudes cívicas de los estudiantes, se utilizó el modelamiento multinivel a través de la técnica de Modelos Jerárquico Lineales (Hierarchical Linear Models). El modelamiento multinivel es una técnica estadística que permite predecir los valores de una variable de interés (variable respuesta), la cual comúnmente es de carácter continuo, dados los valores de una o más medidas relacionadas que son los factores contextuales (Willms, 2006).

Se emplearon cuatro modelos en los que se analizaron los efectos que el estrato educativo y el nivel socioeconómico tienen sobre las actitudes cívicas de los alumnos, y la cantidad de varianza que explican. El primer modelo, denominado nulo o vacío, permitió determinar los montos de varianza en las actitudes cívicas que se asocian a las diferencias entre los alumnos y las escuelas; es el modelo contra el que se comparan los modelos que incluyen las variables estrato escolar y nivel socioeconómico. El segundo modelo incluye la variable de estrato escolar (tipo de secundaria) en el cual se analizó la influencia de los diferentes estratos en las actitudes cívicas de los alumnos, utilizando como valor de referencia las secundarias generales. Este modelo permite conocer la posición que tiene el estrato escolar del resto de las escuelas con respecto a las secundarias generales. En el tercer modelo se introdujo la variable compuesta de nivel socioeconómico de los alumnos y se evaluó la influencia que ésta tiene en las actitudes cívicas de los estudiantes sin considerar el estrato escolar al que pertenecen. Finalmente, el cuarto modelo incluyó el nivel socioeconómico de los alumnos y permitió comparar, simultáneamente, el impacto del estrato escolar y del nivel socioeconómico sobre las actitudes cívicas.

De cada uno de los modelos, para las diferentes escalas que integran el CIVED, se estimó el porcentaje de varianza entre las escuelas y al interior de ellas; analizando la relación existente entre el estrato escolar, el nivel socioeconómico y las actitudes cívicas de los alumnos. Finalmente, se evaluó la relación entre el estrato escolar y las actitudes cívicas cuando se considera el nivel socioeconómico de los estudiantes.

La comparación entre los modelos se realiza mediante el empleo de la devianza, la cual puede ser interpretada como una medida de falta de ajuste entre el modelo y los datos, en donde valores más altos indican un menor ajuste del modelo a los datos; por lo que el modelo que registre una menor devianza será el que mejor explique la variabilidad en las actitudes cívicas.

Por su parte, la metodología para el escalamiento de la variable nivel socio-económico se apoyó en el empleo de la metodología de la Teoría de Respuesta al Ítem, basada en las correlaciones significativas entre los reactivos que integran la variable construida, en la contribución significativa de cada uno de ellos a la variable, en que la bondad de ajuste de los reactivos se encuentren de valores específicos (Linacre, 1998), y que la media de la variable sea cero con una desviación estándar de una unidad (Backhoff *et al.*, 2008; Linacre, 1998).

1.2. Resultados

Al comparar los puntajes registrados en las trece escalas del CIVED con respecto al tipo de secundaria, se observa que las secundarias generales y técnicas obtuvieron, en lo general, puntajes por debajo de la media estatal. Por otro lado, los estratos de secundarias privadas y telesecundarias consistentemente poseen puntajes en las escalas que se encuentran por encima de la media estatal (Tabla 2).

TABLA 2. MEDIAS Y ERRORES ESTÁNDAR DE LAS ESCALAS DEL CIVED POR TIPO DE SECUNDARIA

Escala	Tipo de Secundaria								Baja California	
	General		Técnica		Privada		Telesecundaria		Media	DE
	Media	DE	Media	DE	Media	DE	Media	DE		
Democracia	9.95	1.94	9.99	2.06	10.35	2.07	10.08	1.85	10.01	1.99
Ciudadanía	9.92	2.00	10.04	2.05	10.15	1.83	10.36	1.88	10.00	2.00
Gobierno	9.95	2.01	10.04	1.99	10.10	2.08	10.16	1.84	10.00	2.00
Confianza en las instituciones	9.99	2.02	9.97	1.96	10.06	2.04	10.21	1.83	10.00	1.99
Nuestro país	9.96	1.98	9.99	2.04	10.11	1.98	10.34	1.89	10.00	2.00
Oportunidades 1	9.96	1.96	10.12	2.05	10.03	2.04	9.70	2.03	10.00	2.00
Oportunidades 2	9.91	1.99	10.03	2.00	10.40	1.91	10.09	2.11	10.00	1.99
Escuela	9.95	1.94	9.97	2.06	10.31	2.10	10.06	1.93	10.00	2.00
Aprendizaje	10.00	1.97	9.88	2.03	10.11	2.02	10.53	1.91	10.00	2.00
Acción política 1	9.97	1.99	9.99	1.99	10.11	2.09	10.15	1.95	10.00	2.00
Acción política 2	9.96	2.01	10.06	1.97	10.05	1.94	10.05	2.10	10.00	1.99
Salón	9.93	2.01	9.93	1.99	10.43	1.91	10.43	1.98	10.00	2.00
Sistema Político	9.95	1.99	9.96	1.99	10.41	1.95	10.01	2.07	10.00	2.00

TABLA 3. PROPIEDADES PSICOMÉTRICAS DE LA VARIABLE COMPUESTA NIVEL SOCIOECONÓMICO

Indicadores		1	2	3	4	5	6	7	8	9	10	11	rbis	Alfa sin ítem	Medida	Ajuste
1	Expectativas educativas												0.32	0.64	-0.01	1.44
2	Reproductor de CD	0.08											0.28	0.65	-2.47	0.93
3	Reproductor DVD	0.04	0.36										0.27	0.65	-2.24	0.94
4	Computadora	0.20	0.26	0.24									0.47	0.63	-0.71	0.86
5	Consola de juego de video	0.07	0.26	0.25	0.27								0.29	0.64	-0.45	1.00
6	Internet	0.19	0.24	0.23	0.63	0.30							0.50	0.63	0.35	0.84
7	Teléfono	0.10	0.21	0.21	0.35	0.24	0.41						0.31	0.65	-1.39	0.96
8	Microondas	0.09	0.25	0.27	0.26	0.24	0.24	0.22					0.26	0.65	-1.89	0.97
9	Automóvil	0.05	0.19	0.21	0.28	0.22	0.23	0.22	0.24				0.23	0.65	-2.22	0.99
10	Número de libros	0.24	0.15	0.14	0.30	0.16	0.32	0.22	0.11	0.13			0.40	0.62	0.34	1.14
11	Escolaridad Padre	0.25	0.17	0.17	0.29	0.19	0.35	0.19	0.15	0.14	0.27		0.56	0.58	0.29	0.95
12	Escolaridad Madre	0.27	0.15	0.16	0.29	0.18	0.32	0.16	0.16	0.13	0.29	0.57	0.57	0.57	0.29	0.94

Nota: $p < .01$ (en negritas).

Como paso previo a la inclusión del nivel socioeconómico en los análisis comprometidos, se obtuvieron las propiedades psicométricas de esta variable compuesta. Así, la variable nivel socio-económico quedó

conformada por 12 indicadores los cuales registran correlaciones significativas entre sí con valores del orden de 0.04 a 0.57, un índice de consistencia interna de .65, y correlaciones individuales de cada indicador con la variable general oscilando en un rango de 0.26 a 0.57 (Tabla 3). A pesar de que algunos indicadores no contribuyen de manera significativa a incrementar la consistencia interna de la variable, se decidió conservarlos debido a que la variabilidad que aportan en las medidas de lógitos y sus valores de ajuste se encuentran dentro de los límites deseados (Linacre, 1998).

1.2.1. Modelo Nulo

En la tabla 4 puede apreciarse las estimaciones de los coeficientes y los errores estándar para las diferentes escalas el monto total de varianza explicada por el estudiante y por la escuela, así como el respectivo porcentaje que representa en cada nivel. Para fines de los análisis posteriores deja de ser necesaria la utilización de la escala estandarizada empleada con anterioridad ya los modelos multinivel realizan ajustes y ponderaciones para calcular las medias de cada escuela.

Los resultados de los componentes de la varianza indican que hay pocas variaciones entre escuelas, encontrándose el mayor porcentaje de varianza explicada al nivel de los estudiantes. Las escalas que explican una mayor proporción de varianza al nivel de la escuela son *Democracia* (2.8%), *Gobierno* (2.9%), *Oportunidades 2* (2.0%), *Escuela* (3.3%), *Aprendizaje en la Escuela* (1.8%), y *Salón de Clases* (5.4%). En el resto de las escalas el porcentaje de varianza explicada al nivel de la escuela es marginal, registrando valores por debajo de un punto porcentual.

**TABLA 4. COEFICIENTES Y ERRORES ESTÁNDAR DEL ANÁLISIS MULTINIVEL PARA LAS ESCALAS DEL CIVED:
VARIANZA AL NIVEL DE LOS ESTUDIANTES Y DE LA ESCUELA**

	Escala:												
	Democracia	Ciudadanía	Gobierno	Confianza en Instituciones	Nuestro País	Oportunidades 1	Oportunidades 2	Escuela	Aprendizaje en la Escuela	Acción Política 1	Acción Política 2	Salón de Clases	Sistema Político
Coeficientes	Estimado (EE)	Estimado (EE)	Estimado (EE)	Estimado (EE)	Estimado (EE)	Estimado (EE)	Estimado (EE)	Estimado (EE)	Estimado (EE)	Estimado (EE)	Estimado (EE)	Estimado (EE)	Estimado (EE)
Intercepto	9.99 (0.07)	10.01 (0.06)	10.00 (0.07)	10.03 (0.05)	10.02 (0.05)	10.01 (0.04)	10.04 (0.06)	10.05 (0.07)	10.06 (0.06)	10.03 (0.04)	10.01 (0.04)	10.12 (0.08)	10.04 (0.05)
NSE													
NSE escuela													
Técnica													
Privada													
Telesecundaria													
Componentes de la varianza													
Estudiante	3.92	3.95	3.96	3.96	3.96	3.98	3.94	3.88	3.94	3.98	3.98	3.85	3.96
Escuela	0.11	0.07	0.12	0.06	0.06	0.03	0.08	0.13	0.07	0.02	0.02	0.22	0.05
Porcentaje de varianza													
Estudiante	97.22	98.31	97.09	98.57	98.53	99.30	97.97	96.68	98.18	99.38	99.46	94.61	98.84
Escuela	2.78	1.69	2.91	1.43	1.47	0.70	2.03	3.32	1.82	0.62	0.54	5.39	1.16
Devianza	17329.84	17352.62	17370.75	17355.88	17358.86	17362.60	17344.33	17299.62	17339.12	17362.52	17363.69	17274.22	17354.53

NOTA: Los coeficientes significativos son indicados con negritas.

1.2.2. Modelo estrato escolar

Cuando se controlan los efectos del tipo de escuela, teniendo como valor referencial las secundarias generales, el porcentaje de varianza explicada al nivel de la escuela disminuye en la mayor parte de las escalas; no obstante, los mayores porcentajes de varianza explicados en las escalas prevalecen como en el modelo nulo: *Democracia* (2.3%), *Gobierno* (3.1%), *Oportunidades 2* (1.48%), *Escuela* (3.2%), *Aprendizaje Escuela* (1.5%), y *Salón de Clases* (4.3%) (Tabla 5). El tipo de secundaria tiene un efecto significativo en algunas de las escalas para el estrato escolar *Privada*, la cual presenta coeficientes positivos para siete de las trece escalas. Por su parte el tipo de escuela registra un coeficiente significativo en la escala de *Telesecundaria* sólo tiene un coeficiente significativo en la escala de *Aprendizaje Escuela*.

Considerando que el valor referencial del tipo de secundaria son las generales, los coeficientes deben ser interpretados como las posiciones relativas de cada tipo de secundaria respecto al valor referencial. Así, los estudiantes que asisten a las escuelas secundarias técnicas no difieren significativamente de los de escuelas generales en alguna de las escalas. Por su parte, los estudiantes de *Telesecundaria* se encuentran por encima de los alumnos de secundarias generales en una escala; finalmente, los alumnos que asisten a secundarias privadas muestran mayores puntajes en actitudes cívicas en siete escalas.

En lo que respecta al ajuste de los modelos para cada escala, puede observarse que de acuerdo a los valores de la devianza, la inclusión del estrato escolar no favorece de manera general a los modelos, debido a que incrementa el porcentaje de varianza entre escuelas para seis escalas (*Gobierno*, *Confianza en Instituciones*, *Nuestro País*, *Oportunidades 1*, *Acción Política 1* y *Acción Política 2*); en esas escalas el valor de la devianza fue mayor o igual al obtenido por el modelo nulo.

TABLA 5. COEFICIENTES Y ERRORES ESTÁNDAR DEL ANÁLISIS MULTINIVEL PARA LAS ESCALAS DEL CIVED;
CONTROLANDO LOS EFECTOS DEL TIPO DE ESCUELA

	Escala:												
	Democracia	Ciudadanía	Gobierno	Confianza en Instituciones	Nuestro País	Oportunidades 1	Oportunidades 2	Escuela	Aprendizaje en la Escuela	Acción Política 1	Acción Política 2	Salón de Clases	Sistema Político
Coefficientes	Estimado (EE)	Estimado (EE)	Estimado (EE)	Estimado (EE)	Estimado (EE)	Estimado (EE)	Estimado (EE)	Estimado (EE)	Estimado (EE)	Estimado (EE)	Estimado (EE)	Estimado (EE)	Estimado (EE)
Intercepto	9.90 (0.07)	9.91 (0.06)	9.95 (0.08)	9.99 (0.06)	9.96 (0.06)	10.00 (0.05)	9.91 (0.05)	9.96 (0.06)	10.00 (0.06)	10.01 (0.06)	9.98 (0.05)	9.95 (0.09)	9.96 (0.05)
NSE escuela	-0.08 (0.21)	0.04 (0.17)	-0.09 (0.20)	-0.01 (0.14)	-0.02 (0.16)	0.13 (0.10)	0.05 (0.18)	-0.07 (0.25)	-0.10 (0.16)	-0.00 (0.10)	0.04 (0.11)	-0.07 (0.21)	-0.02 (0.14)
Técnica	0.48 (0.13)	0.25 (0.10)	0.13 (0.18)	0.11 (0.16)	0.23 (0.11)	0.03 (0.13)	0.51 (0.12)	0.36 (0.13)	0.17 (0.11)	0.08 (0.14)	0.13 (0.12)	0.50 (0.19)	0.44 (0.09)
Privada	0.12 (0.12)	0.35 (0.20)	0.27 (0.15)	0.19 (0.18)	0.22 (0.20)	-0.27 (0.18)	0.18 (0.10)	0.16 (0.15)	0.49 (0.19)	0.15 (0.10)	-0.04 (0.15)	0.42 (0.24)	0.11 (0.17)
Telesecundaria													
Componentes de la varianza													
Estudiante	3.92 0.09	3.95 0.06	3.95 0.13	3.96 0.06	3.96 0.06	3.97 0.03	3.94 0.06	3.88 0.13	3.93 0.06	3.98 0.03	3.98 0.02	3.85 0.17	3.96 0.03
Escuela	97.68 2.32	98.47 1.53	96.84 3.16	98.38 1.62	98.52 1.48	99.36 0.64	98.52 1.48	96.82 3.18	98.48 1.52	99.33 0.67	99.42 0.58	95.67 4.33	99.20 0.80
Porcentaje de varianza													
Devianza	17322.83	17350.91	17370.81	17358.37	17359.12	17362.28	17336.47	17297.07	17333.20	17366.25	17367.46	17267.00	17347.08

NOTA: Los coeficientes significativos son indicados con negritas.

1.2.3. Modelo nivel socioeconómico

En este modelo se incorpora la variable compuesta de *Nivel Socioeconómico* sin considerar el estrato escolar, a fin de evaluar sus efectos individuales en las actitudes cívicas de los estudiantes.

Con la inclusión del *Nivel Socioeconómico* de los alumnos hay una disminución significativa en el porcentaje de varianza al nivel de la escuela en 11 de las 13 escalas, en donde la mayor reducción de las varianzas ocurrió en las escalas de *Democracia* (1.0%), *Gobierno* (1.8%), *Oportunidades 2* (0.6%), y *Escuela* (1.9%). En las escalas restantes se observa una reducción en la varianza al nivel de la escuela de menos de un punto porcentual (Tabla 6).

La variable compuesta *Nivel Socioeconómico* se incluyó en el modelo a nivel individual y como un agregado por escuela, resultante de promediar el nivel socioeconómico de los alumnos de una escuela particular. Esta variable agregada de *Nivel Socioeconómico* por escuela permite estimar los efectos contextuales que el promedio de nivel socioeconómico de la escuela tiene en las actitudes cívicas de los alumnos. Los coeficientes para el *Nivel Socioeconómico* de los alumnos resultaron estadísticamente significativos y positivos para 12 de las 13 escalas que integran el CIVED, con valores que oscilan entre 0.16 a 0.30. Por su parte el agregado de *Nivel Socioeconómico* al nivel de la escuela solamente fue significativo, y negativo, en la escala de *Acción Política 1* (-0.26).

El ajuste de los modelos a la varianza sufre cambios significativos en referencia al modelo de estrato escolar, donde las escalas *Democracia*, *Gobierno*, *Confianza en Instituciones*, *Acción Política 1* y *Acción Política 2* registraron una reducción de la devianza. En el resto de las escalas el valor de la devianza fue igual o mayor al obtenido en el modelo que representó previamente.

TABLA 6. COEFICIENTES Y ERRORES ESTÁNDAR DEL ANÁLISIS MULTINIVEL PARA LAS ESCALAS DEL CIVED;
CONTROLANDO LOS EFECTOS DEL NIVEL SOCIOECONÓMICO

	Escala:												
	Democracia	Ciudadanía	Gobierno	Confianza en Instituciones	Nuestro País	Oportunidades 1	Oportunidades 2	Escuela	Aprendizaje en la Escuela	Acción Política 1	Acción Política 2	Salón de Clases	Sistema Político
Coefficientes	Estimado (EE)	Estimado (EE)	Estimado (EE)	Estimado (EE)	Estimado (EE)	Estimado (EE)	Estimado (EE)	Estimado (EE)	Estimado (EE)	Estimado (EE)	Estimado (EE)	Estimado (EE)	Estimado (EE)
Intercepto	10.00 (0.05)	10.01 (0.04)	10.00 (0.06)	10.03 (0.05)	10.02 (0.05)	10.02 (0.04)	10.03 (0.04)	10.04 (0.06)	10.06 (0.05)	10.04 (0.04)	10.01 (0.04)	10.11 (0.07)	10.02 (0.04)
NSE	0.30 (0.05)	0.26 (0.04)	0.22 (0.04)	0.20 (0.04)	0.22 (0.04)	0.05 (0.03)	0.25 (0.04)	0.26 (0.04)	0.18 (0.04)	0.26 (0.04)	0.17 (0.04)	0.16 (0.04)	0.28 (0.04)
NSE escuela	0.09 (0.09)	-0.05 (0.08)	0.00 (0.09)	-0.05 (0.08)	-0.06 (0.09)	0.07 (0.08)	0.10 (0.08)	0.08 (0.10)	-0.12 (0.09)	-0.26 (0.09)	-0.11 (0.08)	0.10 (0.12)	0.04 (0.06)
Técnica													
Privada													
Telesecundaria													
Componentes de la varianza													
Estudiante	3.79	3.87	3.89	3.90	3.90	3.97	3.87	3.82	3.89	3.91	3.94	3.81	3.88
Escuela	0.04	0.03	0.07	0.04	0.04	0.03	0.02	0.07	0.07	0.03	0.02	0.17	0.02
Porcentaje de varianza													
Estudiante	99.02	99.27	98.17	99.00	99.10	99.31	99.42	98.09	98.33	99.36	99.50	95.75	99.60
Escuela	0.98	0.73	1.83	1.00	0.90	0.69	0.58	1.91	1.67	0.64	0.50	4.25	0.40
Devianza	17162.37	17245.82	17280.27	17282.51	17278.46	17344.14	17237.58	17206.79	17282.36	17280.93	17309.36	17215.96	17248.38

NOTA: Los coeficientes significativos son indicados con negritas.

1.2.4. Modelo estrato escolar y nivel socio-económico

En este cuarto modelo de análisis jerárquico multinivel, además de controlar los efectos del tipo de secundaria se controlaron los efectos que el nivel socioeconómico de los alumnos tiene sobre las distintas escalas del CIVED (Tabla 7).

El nivel socioeconómico tiene efectos significativos y positivos en doce de las trece escalas del CIVED, con coeficientes que oscilan entre 0.16 y .30. En lo referente al nivel socioeconómico por escuela, solamente se encontraron efectos significativos positivos en las escalas de *Ciudadanía* (0.11) y *Gobierno* (0.31), mientras que en la escala de *Acción Política 1* el impacto fue negativo (-0.29).

Respecto a los efectos que el tipo de escuela tiene sobre las distintas escalas se encontró que en promedio las secundarias técnicas se encuentran ligeramente por debajo de las generales en las escalas del CIVED, sin embargo, los coeficientes obtenidos no resultaron estadísticamente significativos. Por su parte, las escuelas privadas no difirieron significativamente de las secundarias generales, exceptuando la escala *Gobierno* en donde el coeficiente obtenido fue negativo (-0.52). Finalmente, en las telesecundarias se obtuvieron coeficientes significativos y positivos para siete escalas cuyos valores van de 0.29 a 0.54.

**TABLA 7. COEFICIENTES Y ERRORES ESTÁNDAR DEL ANÁLISIS MULTINIVEL PARA LAS ESCALAS DEL CIVED;
CONTROLANDO LOS EFECTOS DEL TIPO DE ESCUELA Y EL NIVEL SOCIOECONÓMICO**

	Escala:												
	Democracia	Ciudadanía	Gobierno	Confianza en Instituciones	Nuestro País	Oportunidades 1	Oportunidades 2	Escuela	Aprendizaje en la Escuela	Acción Política 1	Acción Política 2	Salón de Clases	Sistema Político
Coeficientes	Estimado (EE)	Estimado (EE)	Estimado (EE)	Estimado (EE)	Estimado (EE)	Estimado (EE)	Estimado (EE)	Estimado (EE)	Estimado (EE)	Estimado (EE)	Estimado (EE)	Estimado (EE)	Estimado (EE)
Intercepto	10.02 (0.06)	9.99 (0.05)	10.07 (0.07)	10.06 (0.05)	10.01 (0.05)	10.04 (0.06)	9.99 (0.05)	10.08 (0.08)	10.03 (0.06)	10.00 (0.07)	9.99 (0.05)	10.02 (0.05)	10.03 (0.05)
NSE	0.30 (0.05)	0.26 (0.04)	0.22 (0.04)	0.20 (0.04)	0.22 (0.04)	0.05 (0.03)	0.25 (0.04)	0.26 (0.04)	0.18 (0.04)	0.27 (0.04)	0.17 (0.04)	0.16 (0.04)	0.29 (0.04)
NSE escuela	0.19 (0.10)	0.11 (0.10)	0.31 (0.14)	0.13 (0.11)	0.01 (0.14)	0.11 (0.08)	0.13 (0.13)	0.26 (0.16)	-0.04 (0.12)	-0.29 (0.11)	-0.16 (0.12)	0.14 (0.18)	0.07 (0.09)
Técnica	-0.11 (0.14)	0.01 (0.11)	-0.10 (0.14)	-0.06 (0.14)	-0.04 (0.13)	0.10 (0.10)	0.02 (0.12)	-0.11 (0.18)	-0.12 (0.16)	0.02 (0.11)	0.05 (0.11)	-0.09 (0.19)	-0.08 (0.12)
Privada	-0.13 (0.18)	-0.19 (0.14)	-0.52 (0.24)	-0.29 (0.21)	-0.06 (0.17)	-0.16 (0.14)	0.05 (0.16)	-0.28 (0.22)	-0.01 (0.17)	0.11 (0.18)	0.11 (0.17)	0.13 (0.30)	0.01 (0.13)
Telesecundaria	0.29 (0.13)	0.50 (0.21)	0.48 (0.15)	0.30 (0.17)	0.32 (0.23)	-0.21 (0.17)	0.32 (0.09)	0.35 (0.15)	0.54 (0.21)	0.14 (0.10)	-0.03 (0.17)	0.53 (0.26)	0.26 (0.13)
Componentes de la varianza													
Estudiante	3.79	3.87	3.89	3.90	3.90	3.97	3.87	3.82	3.89	3.91	3.94	3.81	3.88
Escuela	0.04	0.02	0.06	0.04	0.04	0.03	0.02	0.07	0.05	0.03	0.02	0.15	0.02
Porcentaje de varianza													
Estudiante	98.91	99.56	98.52	99.07	99.08	99.29	99.38	98.09	98.62	99.31	99.41	96.18	99.59
Escuela	1.09	0.44	1.48	0.93	0.92	0.71	0.62	1.91	1.38	0.69	0.59	3.82	0.41
Devianza	17161.83	17240.09	17271.69	17281.17	17278.63	17343.42	17237.98	17204.53	17274.99	17283.91	17313.15	17210.50	17249.07

NOTA: Los coeficientes significativos son indicados con negritas; asimismo, se omiten los errores estándar de los interceptos.

En lo que refiere al porcentaje de varianza explicada se observa una disminución del porcentaje correspondiente a las escuelas en todas las escalas, reduciendo la proporción de varianza explicada por

las escalas, siendo las escalas *Gobierno* (1.5%), *Escuela* (1.9%), *Aprendizaje en la Escuela* (1.4%) y *Salón de Clases* (3.8%) las únicas que lo hacen con valores por arriba del punto porcentual.

Las devianzas de los modelos para las diferentes escalas son menores a las de los previos lo que sugiere que la inclusión del nivel socioeconómico y el estrato escolar al modelo contribuye de manera significativa a explicar la variabilidad en las actitudes cívicas.

2. DISCUSIÓN

A pesar de que los resultados de la presente aproximación sugieren un efecto reducido de la escuela en la explicación de la conformación de las opiniones y actitudes cívicas en los estudiantes de educación secundaria, la importancia que éstos atribuyen a la escuela en su formación ciudadana, y la valoración que hacen del salón de clases como espacio en el que se discuten diversos temas que estimulan el desarrollo de actitudes de naturaleza cívica, son componentes que advierten sobre el papel que juega la escuela en las construcciones que los estudiantes hacen sobre su papel en la sociedad y sobre la forma de concebir al resto de los actores sociales. Estudios realizados en el ámbito internacional han documentado cómo el nivel de involucración de un estudiante en la dinámica y actividades propias de la escuela impacta en sus conocimientos y actitudes de naturaleza cívica (Geske, 2004; Schagen, 2004). Así, el que las prácticas democráticas al interior de la escuela tengan un impacto reducido en la vida de un estudiante, pudiera explicarse por los efectos de enajenación o ideologización implícitos en el curriculum escolar que obstaculizan el sentido de participación social (Torney-Purta, Schwille y Amadeo, 1999).

Por otro lado, la relación observada entre el nivel socioeconómico y las actitudes cívicas es similar a la observada entre éste y dominios curriculares tales como el logro en matemáticas o español. Los resultados de diversas evaluaciones realizadas en este país con estudiantes de secundaria, revelan que pertenecer a una escuela privada, registrar niveles altos en las expectativas de los padres y en la escolaridad de la madre y del padre, así como el ingreso familiar, explican los resultados registrados en dichos dominios curriculares (Backhoff *et al.*, 2006). No obstante, los resultados de dichos estudios, atribuyen mayor impacto a la escuela en los resultados educativos de sus estudiantes que el efecto observado en el caso de las actitudes cívicas.

Ahora bien, el hecho de que la reducción en la variación entre escuelas sea muy poco explicada por el nivel socioeconómico del alumno, contradice lo escrito en la literatura nacional e internacional. Lo anterior podría atribuirse a la ausencia de variabilidad entre los alumnos de la muestra, con pocos estudiantes provenientes de niveles socioeconómicos alto y bajo, situación documentada en otros estudios realizados en esta entidad (Carvallo, Caso y Contreras, 2007). Otra posible explicación es la suma de condiciones del contexto fronterizo local que requiere ser interpretado de manera reservada, pues en la franja de la frontera con Estados Unidos, particularmente la de Baja California con California, las posesiones de bienes materiales, tales como el reproductor de discos compactos, DVD, automóvil, horno de microondas, entre otros, difícilmente podrían considerarse indicadores adecuados del nivel socioeconómico. En esta parte, el argumento explicativo expresa la posibilidad de que la ecuación propuesta para determinar el índice de nivel socioeconómico no sea la más precisa ni la más plausible para esta entidad.

Los resultados de este estudio registran ligeras variaciones en el porcentaje de varianza explicada dependiendo del tipo de escuela del que se trata. Sin embargo no se aprecian diferencias notables en dichas variaciones sugiriendo que éstas se encuentran asociadas en mayor medida con las características individuales de los estudiantes que a las de las escuelas. Así, los estudiantes de escuelas privadas obtuvieron en promedio puntajes más elevados de lo que lo hicieron los alumnos de secundarias generales, resultados matizados por el tipo de alumnos que asisten a las mismas, y por sus antecedentes sociales, culturales y económicos, aspectos responsables de las diferencias en dichos puntajes.

Finalmente, al controlar tanto los efectos del tipo de escuela como del nivel socioeconómico, las variaciones en los coeficientes de las escuelas sugieren que hay efectos del nivel socioeconómico que influyen en los puntajes del CIVED. Este supuesto es apoyado por el comportamiento de los coeficientes de las escuelas privadas y de las telesecundarias: mientras que para las primeras los coeficientes disminuyeron, para las segundas se obtuvieron un mayor número de coeficientes significativos de los resultantes en el modelo en el que sólo se controlaron los efectos del tipo de escuela, confirmando diferencias económicas y sociales entre las escuelas.

Los alcances de la presente aproximación se habrán de sumar al conjunto de estudios que recientemente se vienen registrando en México como parte del movimiento de investigación de eficacia escolar. Con lo anterior no sólo se sientan las bases para la elaboración de modelos comprensivos y globales que expliquen la realidad del sistema educativo nacional y que ayuden a orientar las iniciativas para su transformación, sino se contribuye a ampliar nuestro conocimiento sobre las características específicas del entorno local y a identificar las acciones requeridas en materia de planeación y aplicación de programas orientados a la formación cívica y valoral de los estudiantes.

REFERENCIAS BIBLIOGRÁFICAS

- Backhoff, E., Andrade, E., Peón, M., Sánchez, A. y Bouzas, A. (2006). *El aprendizaje del español, las matemáticas y la expresión escrita en la educación básica en México: Sexto de primaria y tercero de secundaria. Resumen Ejecutivo*. México: INEE.
- Backhoff, E., Bouzas, A., Hernández, E. y García, M. (2007). *Aprendizaje y desigualdad social en México. Implicaciones de política educativa en el nivel básico*. México: INEE.
- Backhoff, E., Bouzas, A., Contreras, C., Hernández, E. y García, M. (2007). *Factores escolares y aprendizaje en México. El caso de la educación básica*. México: INEE.
- Backhoff, E., Andrade, E., Sánchez, A. y Peón, M. (2007). *El aprendizaje en tercero de primaria en México: Español, Matemáticas, Ciencias Naturales y Ciencias Sociales*. México: INEE.
- Backhoff, E., Andrade, E., Sánchez, A. y Peón, M. (2008). *El aprendizaje en tercero de preescolar en México*. México: INEE.
- Carvalho, M., Caso, J. y Contreras, L. A. (2007). Estimación del efecto de variables contextuales en el logro académico de estudiantes de Baja California. *Revista Electrónica de Investigación Educativa*, 9 (2). Consultado el día 23 de febrero de 2009 en: <http://redie.uabc.mx/vol9no2/contenido-carvalho.html>
- Caso, J., González_Montesinos, J. M., Rodríguez, J. C. y Urias, E. (2008). *Evaluación de la Educación Cívica en estudiantes de secundaria en Baja California*. UEE Reporte Técnico 08-001. Ensenada: Universidad Autónoma de Baja California.

- Coleman, J.S., Campbell, E.Q., Hobson, C.J., McPartland, J., Mood, A.M., Weinfeld, F.D. y York, R.L. (1966). *Equality of Educational Ppportunity*. Washington: US Government Printing Office.
- Duru-Bellat, M. (2004). *Social inequality at school and educational policies. Fundamentals of Educational Planning series, No.78*. Paris: UNESCO.
- Geske, A. (2004). *Influence of contextual factors on student achievements in mathematics, science and civic knowledge in Latvia*. Disponible en la dirección electrónica: http://www.iea.nl/fileadmin/user_upload/IRC2004/Geske.pdf
- Hansen K. Y., Rosén, M. y Gustafsson, J. E. (2001). *Effects of socio-economic status on reading achievement at collective and individual levels in Sweden in 1991 and 2001*. Disponible en la dirección electrónica: http://www.iea.nl/fileadmin/ user_upload/IRC2004/Hansen_Rosen_Gustafsson1.pdf
- Hanushek, E. A. y Luque, J. A. (2003). Efficiency and equity in schools around the world. *Economics of Education Review*, 22(5), pp. 481-502.
- Inglehart, R. y Baker, W. (2000). Modernization, cultural change and the persistence of traditional values. *American Sociological Review*, 65, pp. 19-51.
- Lagos, M. (2005). The Role of western world power in promoting democracy and defeding peace: Latin American and European Perception. *International Journal of Public Opinion Research*, 17(2).
- Linacre, J. M. (1998). Detecting multidimensionality: which residual data-type work best? *Journal of Outcome Measurement*, 2(3), pp. 266-183.
- Murillo, F.J. (2003). El movimiento de investigación de Eficacia Escolar. En F.J. Murillo (Coord.), *La investigación sobre Eficacia Escolar en Iberoamérica. Revisión internacional del estado de la cuestión*. Bogotá: Convenio Andrés Bello.
- Rodríguez, J. C., Caso, J., Contreras, L. A. y Urias, E. (2007). Cuestionario de Contexto. Universidad Autónoma de Baja California.
- Schagen, I. (2004). Further statistical analysis of the IEA civics data to investigate attitudes to citizenship in England.
- Torney-Purta, J., Schwille, J. y Amadeo, J. A. (1999). *Civic Education Across Countries: Twenty-four National Case Studies from the IEA Civic Education Project*. Amsterdam: IEA.
- Torney-Purta, J. y Henry-Barber, J. (2004). *Democratic school participation and Civic Attitudes among European adolescents: Analysis of data from the IEA Civic Education Study*. Council of Europe: Education for Democratic Citizenship 2001-2004.
- Valenti, G. (2007) *Factores asociados al logro educativo de matemáticas y español en la Prueba ENLACE 2007*. México: SEP.
- Willms, J. D. (2006). *Learning divides: Ten policy questions about the performance and equity of schools and schooling Systems*. Paris: UNESCO.
- Willms, J. D. (2003). Ten hypotheses about socioeconomic gradients and community differences in children's developmental outcomes. Informe preparado para el Human Resources Development Canada. Disponible en la dirección electrónica: <https://auxweb.unb.ca/applications/crisp/pdf/0305.pdf>
- Willms, J. D. (2001). Monitoring school performance for "standards-based reform". *Evaluation and Research in Education*, 14(3-4), pp. 237-253.

DA AVALIAÇÃO DE PROFESSORES: REFLEXÕES SOBRE O CASO PORTUGUÊS

Maria Assunção Flores

Revista Iberoamericana de Evaluación Educativa 2009 - Volumen 2, Número 1

<http://www.rinace.net/riee/numeros/vol2-num1/art13.pdf>

Fecha de recepción: 13 de marzo de 2009

Fecha de comunicación de dictamen: 22 de abril de 2009

Fecha de aceptación: 22 de abril de 2009

"Sou um professor com o ensino no coração e há momentos na sala de aula em que mal posso conter a minha alegria. Quando eu e os meus alunos descobrimos novos territórios inexplorados, quando a bruma que invadia o caminho se dissipa diante de nós, quando a nossa experiência é iluminada pela vida brilhante da mente – nesses instantes, ensinar é a melhor profissão do mundo. Contudo, noutros momentos, a sala de aula parece estar tão desanimada, cheia de dificuldades ou até mesmo confusa – e sinto-me impotente para mudar esta situação – que a minha pretensão de ser professor parece tornar-se num mero fingimento transparente. Nestes momentos, vejo o inimigo em todo o lado: nos alunos que parecem vir de um planeta longínquo, naquela disciplina que eu pensava conhecer e na patologia pessoal que me faz continuar a ganhar a minha vida desta forma (...). Na nossa pressa de reformar a educação, esquecemo-nos de uma verdade simples: a reforma nunca será realizada renovando simplesmente as contratações, ou reestruturando as escolas, reescrevendo os currículos, ou revendo textos, se continuarmos a desvalorizar e a desanimar os recursos humanos denominados professores, de quem tanto depende... se não conseguirmos valorizar – e desafiar – o coração humano, que é a fonte do bom ensino".

Parker J. Palmer, *The Courage to Teach* (1998:1-3)

Este excerto de Palmer, retirado da obra "A coragem de ensinar", ilustra bem algumas das características da profissão docente. O entusiasmo, a paixão, o empenho, a dedicação, as recompensas intrínsecas, o lado cognitivo e emocional do ensino são alguns dos aspectos que estão bem patentes neste testemunho, mas também ele expressa, de forma clara, a complexidade, as dificuldades, as exigências e os desafios inerentes à profissão docente. As identidades profissionais dos professores – quem são, a imagem que têm de si próprios, os sentidos que atribuem ao seu trabalho e a si próprios enquanto professores e os significados que os outros lhes atribuem – dependem de uma diversidade de factores: da forma como se relacionam com os seus alunos, da disciplina que ensinam, dos papéis que desempenham, dos contextos em que trabalham, do quadro social e cultural mais amplo em que o ensino se inscreve, da sua biografia pessoal, das oportunidades de aprendizagem e de desenvolvimento profissional no local de trabalho, do apoio, etc.

Enquanto processo culturalmente inscrito e, por isso, aberto e mutável, a identidade inclui determinados sentidos, valores e imagens do que significa ser professor, o que implica a combinação, por vezes conflitual, entre pedaços do passado e a situação presente (Feiman-Nemser, 2003). É, portanto, no dizer de Nóvoa (1992:16), "um lugar de lutas e de conflitos", "um espaço de construção de maneiras de ser e de estar na profissão" e não "um dado adquirido". Que entendimento(s) têm os professores de si próprios, da sua profissão e dos contextos em que exercem a sua actividade? Como se vêem e como vêem os contextos de mudança em que trabalham? Como definem o seu profissionalismo docente face ao enquadramento político, social e cultural em que se inscreve o ensino e às mudanças que têm sido introduzidas? A resposta a estas interrogações é complexa e assenta num conjunto de dimensões inerentes ao trabalho docente e ao modo como ele é perspectivado num dado momento e num determinado contexto. Mas estas questões são fundamentais para compreender o modo como os professores vivem a mudança no seu quotidiano profissional, sobretudo quando ela implica alterações profundas no estatuto da carreira docente e na avaliação do desempenho docente de que falaremos mais adiante.

Numa análise dos contextos de mudança em educação, há autores que aludem à "patologia da mudança "top-down", à "ironia sistemática" e à lógica da recentralização que têm marcado a introdução de alterações no trabalho das escolas e dos professores (ver, por exemplo, A. Hargreaves, 1998), quando se

preconiza a mudança através de procedimentos burocráticos, o que implica um conjunto de “tensões” e “dilemas” no seu processo de implementação. Por outras palavras, a imposição de mudanças por parte da Administração Central assume, de forma implícita, que a sua implementação constitui um acontecimento, um facto e não um processo. Fullan (2001:71) chama a atenção para o conjunto e dinâmica de factores e de “dilemas intrínsecos no processo de mudança” que tornam o seu sucesso num “processo subtil e complexo”. Este processo é ainda mais complexo quando se trata da questão da avaliação de professores e do estatuto da carreira docente, a que subjaz uma determinada visão de professor e de ensino.

1. CONCEPÇÕES DE PROFESSOR E DE PROFISSIONALISMO DOCENTE

Day (1999:12) traduz, de forma clara, a co-existência de diferentes discursos, por vezes conflituais, que traduzem distintos modos de encarar o profissionalismo docente quando afirma:

“...as alterações na prática operacional do “profissionalismo” reflectem as crescentes complexidades e contradições inerentes ao trabalho dos professores num mundo pós-moderno (...) o seu trabalho encerra ao mesmo tempo um desafio e uma ameaça. Tanto podem ser autónomos, como responsáveis perante outros, independentes como colaboradores, controlar o seu trabalho e não o controlar, centrados no professor como centrados no aluno”.

Estudos empíricos apontam para posições de alguma ambivalência relativamente ao modo como os professores encaram, por exemplo, a autonomia curricular (Morgado, 2003) e para um conjunto de ambiguidades e tensões no modo como vêem o seu trabalho (Flores, 2003, 2005). Por outras palavras, os professores podem responder de diferentes modos às mudanças “impostas”, co-existindo, portanto, diferentes visões do profissionalismo docente, ou, nas palavras de Whitty (2000), “diferentes tipos de profissionalismo”. Estas perspectivas encerram diferentes leituras do ensino e do que significa ser professor. Para além do contexto social, político, cultural e económico em que o ensino se situa, o(s) sentido(s) que os próprios professores atribuem ao seu trabalho diário – com as suas potencialidades, constrangimentos e limitações – constituem aspectos-chave na (re)definição da natureza do profissionalismo docente.

Como sustentam Hargreaves e Goodson (1996:22-23),

“...o profissionalismo docente é aquilo que os professores e outros experienciam enquanto tal e não aquilo que os decisores políticos e outros afirmam que deveria ser. A experiência do profissionalismo ou a sua negação tem de ser encontrada no estudo do dia-a-dia do trabalho dos professores.”

A este respeito vale a pena referir alguns resultados de um estudo (Flores, Day e Viana, 2007) que pretendeu analisar e compreender a(s) forma(s) como os professores (portugueses e ingleses) entendiam as mudanças mais recentes ocorridas no ensino, bem como os seus efeitos no modo como encaravam o profissionalismo docente e a(s) sua(s) identidade(s) profissional(ais). Quando questionados sobre quais as dimensões do seu profissionalismo que estavam a ser ameaçadas, a maioria dos participantes no estudo referiu o exercício do seu juízo discricionário, os propósitos morais e sociais, o comprometimento/dedicação e a atenção/preocupação, salientando ainda que as culturas colaborativas, a aprendizagem contínua e a complexidade das tarefas constituíam aspectos que, na sua opinião, estavam a ser promovidos. Dos discursos dos professores que participaram neste estudo emergiu um conjunto de constrangimentos no desenvolvimento do seu trabalho (Flores e Viana, 2004), destacando-

se, entre outros, os seguintes: i) a ausência de uma visão clara em termos de quadro político e legal, o que conduz a atitudes ambíguas e paradoxais; ii) a falta de (in)formação, de comunicação e de recursos para as escolas e para os professores; iii) a cultura profissional marcada pelo isolamento (alguns professores falam de uma “cultura da solidão”); iv) a burocracia, mobilidade docente e a falta de motivação; v) a ausência de entendimento e apropriação da mudança e dos seus conceitos-chave, o que leva à co-existência de dois tipos de discurso, um – retórico - que tende a colar-se às expectativas, exigências e linguagem das políticas curriculares, outro – a um nível prático – que remete para a ausência de mudança. Destes resultados ressalta uma concepção de profissionalismo marcada pela ambivalência, pelo conflito e pela ausência de referentes claros. No entanto, a par destes constrangimentos, foram também identificados alguns aspectos positivos, nomeadamente: i) o reconhecimento de algumas características do profissionalismo docente (culturas colaborativas e aprendizagem contínua que os professores tendem a valorizar); ii) uma perspectiva ecológica do ensino e da aprendizagem (multidimensionalidade do processo de aprendizagem); iii) um entendimento mais amplo dos papéis/funções dos professores e das escolas; iv) um entendimento da escola enquanto contexto organizacional que envolve (ou pode envolver) os seus intervenientes num fim comum (com propósitos e objectivos partilhados); v) uma maior consciencialização do papel do professor enquanto agente de mudança e vi) o reconhecimento da importância da avaliação enquanto factor de motivação e de valorização do profissionalismo docente. E é exactamente sobre esta última questão que iremos reflectir nas secções seguintes deste artigo partindo de uma perspectiva mais global da problemática da avaliação de professores para depois nos centrarmos num conjunto de aspectos mais marcantes da avaliação do desempenho docente no contexto português (actualmente em fase de implementação).

2. O QUE ESTÁ EM CAUSA NA AVALIAÇÃO DE PROFESSORES?

Quando se aborda a questão da avaliação de professores, é necessário reflectir sobre um conjunto de aspectos que a enquadram e definem, nomeadamente a concepção do que significa ser profissional / professor e o modo como se encaram as funções e as dimensões inerentes à profissão docente. Como sustenta Hadji (1994a:32),

“A dificuldade de avaliar os professores deriva muito mais da incerteza que sobreleva a própria essência do ensino e da ausência de consensos a esse respeito, do que de problemas técnicos, sempre subalternos, para não dizer secundários. (...) O objecto da avaliação dos professores é difícil de estabelecer, porque difícil de definir.”

Assim, um sistema de avaliação de professores encontra-se associado a determinadas visões de ensino e de professor (Darling-Hammond, Wise e Pease, 1983; Sachs, 2003; Flores, 2005; Day, Flores e Viana, 2007) e ao modo como essas concepções de ensino são traduzidas em padrões (standards), critérios e indicadores de avaliação (Avalos e Assael, 2006).

2.1. Propósitos da avaliação, qualidade docente e padrões de ensino

A literatura neste domínio tem demonstrado a complexidade (e controvérsia) da avaliação de professores tendo em conta os seus propósitos, processos e efeitos. De um modo geral, tem-se identificado a tensão entre propósitos formativos (orientados para o desenvolvimento profissional) e sumativos (baseados na prestação de contas e orientados para decisões de gestão de carreira) (Day, 1992; Chow et al., 2002; Avalos e Assael, 2006; Stronge, 2006). Neste campo, há quem advogue a incompatibilidade destes dois

propósitos, enquanto outros defendem a possibilidade de incorporar propósitos formativos e sumativos num mesmo sistema de avaliação (sobre este debate, veja-se, por exemplo, Simões, 2000; Chow et al., 2002; Avalos e Assael, 2006).

A esta discussão não é alheia a problemática da definição dos padrões (standards) de ensino que servem como referente à avaliação de professores. A este propósito, Darling-Hammond (2008) clarifica os conceitos de qualidade dos professores e qualidade do ensino, que, embora estejam relacionados, são distintos. A qualidade dos professores diz respeito ao conjunto de traços pessoais, destrezas e perspectivas que alguém traz para o ensino, incluindo também determinadas disposições em termos de comportamento. A autora, com base na investigação sobre a eficácia docente, identifica um conjunto de qualidades, entre as quais se destacam a capacidade de comunicação, o domínio do conhecimento do conteúdo e do conhecimento pedagógico do conteúdo, a compreensão dos aprendentes e da sua aprendizagem e desenvolvimento, ser um "especialista/expert adaptativo" que permite ao professor julgar sobre o modo como deve agir em determinado contexto para responder às necessidades dos alunos. Outros aspectos incluem um conjunto de disposições para apoiar e estimular a aprendizagem dos alunos, para ensinar de modo justo e imparcial, para adaptar o ensino de modo a que os alunos obtenham sucesso, para continuar a aprender ao longo da carreira, etc. A mesma autora refere que estas qualidades estão presentes nos standards adoptados, por exemplo, pelo *National Board for Professional Teaching Standards* (1999) para a certificação dos professores nos EUA. Estes padrões baseiam-se em cinco proposições centrais: i) os professores estão comprometidos com os alunos e a sua aprendizagem; ii) os professores conhecem a disciplina que ensinam e como ensiná-la aos alunos; iii) os professores são responsáveis pela gestão e monitorização da aprendizagem dos alunos; iv) os professores reflectem sistematicamente sobre a sua prática e aprendem pela experiência; v) os professores são membros de comunidades de aprendizagem.

Por seu turno, para Darling-Hammond (2008), a qualidade do ensino surge associada a um bom ensino no sentido de fazer com que os alunos aprendam, em função das exigências da disciplina, dos objectivos pedagógicos e das necessidades dos alunos num determinado contexto. Assim, para esta autora,

"A qualidade do ensino é, em parte, uma função da qualidade do professor – conhecimento, destrezas e disposições do professor – mas é também fortemente influenciada pelo contexto de ensino. (...) Um professor "de qualidade" pode não ser capaz de proporcionar um ensino de qualidade num contexto onde se verifica um desfasamento entre as exigências da situação e o seu conhecimento ou destrezas (...) Por exemplo, um professor preparado para ensinar alunos do ensino secundário pode não ser capaz de ensinar alunos mais novos; um professor que é capaz de ensinar alunos com elevadas destrezas e competências pode não ser capaz de ensinar alunos que têm dificuldades ou que não dispõem em casa dos recursos que o professor está habituado a assumir que existem" (2008:4).

Darling-Hammond (2008) ressalta ainda um outro aspecto da qualidade do ensino: as condições em que ele se realiza. Se um professor de qualidade não possuir materiais curriculares relevantes, equipamentos, uma razoável dimensão das turmas e oportunidades para planificar com outros colegas, etc., a qualidade do ensino pode ser menor, apesar de a qualidade do professor ser elevada. Como refere ainda a autora, "Muitas das condições do ensino estão fora do controle dos professores e dependem dos sistemas político e administrativo em que eles trabalham" (Darling-Hammond, 2008, p.5). E, advogando que a qualidade dos professores pode aumentar a probabilidade de um ensino de qualidade, mas não o garante por si só, acrescenta:

“As iniciativas para desenvolver um ensino de qualidade devem ter em consideração não só o modo de identificar, recompensar e utilizar as destrezas e competências dos professores, mas também o modo como se podem desenvolver contextos de ensino que permitam aos professores realizar boas práticas. (...) Assim, as políticas que constroem os contextos de ensino devem ser ponderadas a par das qualidades e papéis dos professores” (2008:5).

Por seu lado, para Simth (2005), a clarificação de um conjunto de *standards* norteadores da profissão pode constituir um referente importante para a definição ou entendimento do que é um ensino de qualidade e para a sua avaliação. Partindo da literatura neste domínio, a autora faz uma síntese das principais vantagens e desvantagens da utilização dos padrões de ensino. Quanto ao primeiro caso, eles podem, por exemplo, servir como base de entendimento sobre o que os professores devem saber, desenvolver/pensar e fazer; como referente para a avaliação; como meio de comunicar com o público (e outros agentes educativos); como meta ou referente para o desenvolvimento profissional; como base para a reflexão e para o desenvolvimento de um diálogo profissional, etc.

Contudo, como argumenta Smith (2005), os *standards* encerram também um conjunto de questões, nomeadamente a ausência de consenso e o facto de poderem conduzir a uma visão limitada do ensino e da aprendizagem (e condicionar a introdução de novas iniciativas e novas ideias promovidas num quadro de autonomia profissional). Por outro lado, segundo Smith (2005), corre-se o risco de poder ser relegada para plano secundário a importância dos aspectos afectivos do ensino (particularmente numa visão demasiado mercantilista da educação que se centra sobretudo nos resultados). Acresce ainda o facto de o conhecimento que os professores desenvolvem na prática (que é tácito, pessoal, privado, prático, que se baseia na sua experiência, e que inclui valores, perspectivas de encarar o ensino, etc.) não ser facilmente traduzível (se é que é possível) na linguagem dos *standards*. Para além disso, os professores desenvolvem o seu conhecimento e o seu ensino num contexto específico e o que é relevante e apropriado num determinado contexto pode não o ser noutro. Também Day (2004:27-28) identifica alguns aspectos críticos, nomeadamente o facto de “os standards mensuráveis terem em conta um número relativamente reduzido de ensino, aprendizagem e aproveitamento escolar”, o facto de o “uso de competências comportamentais para medir as capacidades dos professores não ter em conta os seus propósitos morais mais amplos” e o facto de “ainda não ter sido provado que a redução da autonomia dos professores tenha originado um melhor ensino ou professores mais comprometidos”.

2.2. Para que serve a avaliação?

A avaliação constitui, portanto, um processo complexo que requer escolhas, decisões, juízos e procedimentos, os quais se inscrevem em modelos que lhe dão sentido e forma, servindo, portanto, múltiplas funções. Falar de avaliação implica, portanto, falar dos pressupostos, das lógicas e dos actores nela implicados. Por outras palavras, implica a reflexão e a explicitação de algumas questões pertinentes: Porquê avaliar? Para quê? Para quem? Para que tipo de decisões? Trata-se, pois, de clarificar as “intencões subjacentes aos projectos de avaliação” (Hadji, 1994b). É evidente que a clarificação destas variáveis de sentido ou axiológicas da avaliação tem implicações na determinação ou na identificação das variáveis mais técnicas que implicam “formas possíveis de avaliação, procedimentos a pôr em prática, operações concretas a realizar” (Hadji, 1994b:46) e que podem traduzir-se nas seguintes interrogações: O que avaliar? Quem avalia? Como? Quando? E, fundamentalmente, quais são os referentes dessa avaliação?

A avaliação do ensino e dos professores pode estar associada a vários propósitos desde a prestação de contas ao desenvolvimento profissional. É necessário, assim, explicitar a seguinte questão: a avaliação serve para quê? Para aumentar a eficácia dos professores? Para promover oportunidades de formação e de desenvolvimento profissional adequadas e significativas? Para melhorar a qualidade do ensino e/ou da aprendizagem? Para melhorar os resultados dos alunos? Para controlar processos e práticas, programas e estratégias? Para medir o desvio entre resultados (objectivos concretizados) e objectivos propostos? Para reestruturar currículos e programas? Para gerir carreiras? Esta multiplicidade de perspectivas – todas elas possíveis e válidas e não totalmente exclusivas, embora, nalguns casos, difíceis de compatibilizar na prática – remete-nos para modos possíveis de encarar a avaliação de professores em função do enfoque e dos seus propósitos e efeitos (previstos e não previstos).

Simth (2005:96) identifica quatro funções principais na avaliação: i) como *gate-keeping* (para determinar a entrada do ensino) que se situa no final da formação inicial, tratando-se, sobretudo de uma avaliação sumativa (orientada para a certificação); ii) para a prestação de contas em relação aos diferentes *stakeholders* (agentes educativos), com base nos resultados dos alunos, estando em causa uma avaliação externa em relação ao professor, de carácter sumativo e com *feedback* limitado; iv) com vista à promoção com efeitos na progressão na carreira, sendo uma avaliação interna e externa, centrando-se no conhecimento e comportamento profissional, tendo um carácter eminentemente sumativo, com vista à tomada de decisão, assumindo, por isso, uma natureza mais formal, mas podendo incorporar uma dimensão formativa através do *feedback* informativo ao professor (no sentido da melhoria da qualidade das suas actividades profissionais); iv) orientada para o desenvolvimento profissional, que assume um carácter formativo, informal, contínuo e sistemático, envolvendo o professor na sua auto-avaliação.

Neste enquadramento, trata-se, pois, de saber quem avalia o quê, para quê e para quem, como e com que finalidade sabendo que a qualidade não é uma questão estática que, uma vez desenvolvida, permanece inalterável (independentemente da fase da carreira, das especificidades dos contextos, etc.). Assim, é fundamental ter em consideração um conjunto de aspectos: a necessidade e a importância de ter em conta o contexto em que a avaliação se concretiza (as suas características e especificidades); a necessidade de tornar explícito o implícito, isto é, o conjunto de propósitos e de intenções subjacentes à avaliação; a necessidade de considerar a multidimensionalidade do ensino – portanto, incluindo todas as vertentes ou dimensões em análise; a participação dos actores – isto é, a avaliação deve reflectir as perspectivas dos diferentes actores envolvidos, por exemplo, não só no que diz respeito aos seus pontos de vista ou posicionamentos sobre o que se avalia, mas também a compreensão e participação na definição dos referentes dessa mesma avaliação. Por outras palavras, trata-se sobretudo de enfatizar uma *perspectiva de implicação* e uma lógica de co-construção de sentido da avaliação. Por fim, é de realçar a importância de considerar a diversidade de fontes, de elementos e de procedimentos de recolha de informação.

3. A AVALIAÇÃO DE DESEMPENHO DOCENTE EM PORTUGAL: ALGUNS PROBLEMAS E DESAFIOS

Em vários países, a preocupação em elevar os padrões de ensino e em melhorar a qualidade das aprendizagens dos alunos tem levado os governos a introduzir reformas nas escolas e no trabalho dos professores no sentido de uma maior prestação de contas, entre as quais se destaca a avaliação dos professores (Middlewood e Cardno, 2001; Avalos, 2004; Avalos e Assael, 2006; Assael e Pavez, 2008). Também em Portugal se introduziu um novo modelo de avaliação de professores cujo processo de

implementação tem sido marcado pela turbulência e contestação, embora, de um modo geral, se reconheça a necessidade de uma tal medida. Este facto pode ser justificado, entre outros factores, pela ineficácia do modelo anterior (baseado essencialmente na auto-avaliação) tendo sido reconhecido na literatura sobre este domínio, no contexto português, como uma avaliação que “não avaliava” (ver, por exemplo, Silva e Conboy, 2004; Pacheco e Flores, 1999).

3.1. A reestruturação da carreira docente

A legislação portuguesa que regulamentou, até há bem pouco tempo, a carreira docente foi publicada em 1990 (Decreto-Lei nº 139-A/90, 28 de Abril). Baseava-se no princípio da “carreira única”, isto é, comum a todos os professores e era composta por 10 escalões, aos quais todos os docentes podiam aceder com base apenas no tempo de serviço e na obtenção de um determinado número de créditos de formação contínua (1 crédito=25 horas de formação). De um modo geral, tratava-se de um modelo meramente administrativo e burocrático que veio legitimar a iniciativa do governo no âmbito da reestruturação da carreira docente: “a avaliação do desempenho, com raras excepções apenas, converteu-se num simples procedimento burocrático, sem qualquer conteúdo”, diz-se no preâmbulo do Decreto-Lei nº 15/2007, de 19 de Janeiro que altera o Estatuto da Carreira Docente.

Em Maio de 2006, surge uma proposta de revisão do estatuto da carreira docente e, com esta, a alteração da avaliação de desempenho dos docentes no sistema educativo português. Com o Decreto-Lei 15/2007, introduz-se o princípio da diferenciação e definem-se mecanismos de avaliação para os professores. A carreira docente surge estruturada em categorias hierarquizadas – “professores titulares” e “professores” – cabendo aos primeiros as funções de coordenação, supervisão e avaliação de outros professores. O acesso ao topo da carreira encontra-se limitado a um terço do número de lugares em cada escola, sendo que o recrutamento para a categoria de professor titular é feito através de concurso mediante as seguintes condições: ter pelo menos 18 anos de serviço docente com avaliação de desempenho igual ou superior a Bom e ser aprovado/a em prova pública que incide “sobre a actividade profissional desenvolvida pelo docente com vista a demonstrar a sua aptidão para o exercício das funções específicas da categoria de professor titular” (artigo 38º, Decreto-Lei nº 15/2007) ¹.

Além disso, o Decreto-Lei nº 15/2007 define as condições de acesso à profissão no sentido de seleccionar os candidatos a professor. Se, até agora, as instituições de ensino superior de formação de professores conferiam uma certificação académica e profissional, única condição de acesso à profissão, passou a ser também exigida aos candidatos a realização de uma “prova de avaliação de conhecimentos e competências” que tem como objectivo avaliar a preparação científica dos candidatos a professor. Introduziu-se ainda o período probatório, que visa “verificar a capacidade de adequação do docente ao perfil de desempenho profissional exigível” tendo a duração mínima de um ano. Durante este período, o professor é “acompanhado e apoiado, no plano didáctico, pedagógico e científico, por um professor titular, detentor, preferencialmente, de formação especializada na área da organização educacional e

¹ Este tem sido um dos aspectos críticos e que tem gerado controvérsia, não só pela hierarquização da carreira que ele encerra, mas pelo facto de, numa fase de transição, o concurso para professor titular sobrevalorizar aspectos mais ligados ao desempenho de cargos (ex. Presidente do Conselho Executivo, membro do Conselho Executivo, Coordenador de departamento, etc.) em detrimento de aspectos ligados ao ensino (a assiduidade foi um dos critérios incluídos). Acresce ainda o facto de, neste concurso, e para efeitos de avaliação da experiência profissional, se ter apenas considerado o período compreendido entre 1999/2000 e 2005/2006 (isto é, 7 anos no total), independentemente do número de anos de serviço que os professores possuíam.

desenvolvimento curricular, supervisão pedagógica e formação de formadores e com menção igual ou superior a *Bom* na última avaliação do desempenho” (artigo 31º, Decreto-Lei nº 15/2007).

É também com o Decreto-Lei nº 15/2007 que se introduzem alterações à avaliação de desempenho docente que deixa de estar baseada na auto-avaliação (através de um relatório de reflexão crítica) para passar a incluir um conjunto de dimensões, nomeadamente a observação de aulas. Justifica-se, assim, um regime de avaliação de desempenho “mais exigente e com efeitos no desenvolvimento da carreira que permita identificar, promover e premiar o mérito e valorizar a actividade lectiva”, diz-se no preâmbulo do referido documento. A avaliação do desempenho visa “a melhoria dos resultados escolares dos alunos e da qualidade das aprendizagens e proporcionar orientações para o desenvolvimento pessoal e profissional no quadro de um sistema de reconhecimento do mérito e da excelência” (artigo 40º do Decreto-Lei nº 15/2007). São ainda objectivos da avaliação do desempenho: “i) contribuir para a melhoria da prática pedagógica do docente; ii) contribuir para a valorização e aperfeiçoamento individual do docente; iii) permitir a inventariação das necessidades de formação do pessoal docente; iv) detectar os factores que influenciam o rendimento profissional do pessoal docente; v) diferenciar e premiar os melhores profissionais; vi) facultar indicadores de gestão em matéria de pessoal docente; vii) promover o trabalho de cooperação entre os docentes, tendo em vista a melhoria dos resultados dos alunos; viii) promover a excelência e a qualidade dos serviços prestados à comunidade” (artigo 40º do Decreto-Lei nº 15/2007). O Decreto Regulamentar nº 2/2008, de 10 de Janeiro, vem regulamentar a avaliação do desempenho docente cujas principais características apresentamos, sumariamente, na secção seguinte.

3.2. Principais aspectos da avaliação do desempenho docente em Portugal

A avaliação do desempenho docente realiza-se com base em quatro dimensões que constituem o perfil geral de desempenho profissional dos educadores de infância e dos professores dos ensinos básico e secundário (Decreto-Lei 240/2001, de 30 de Agosto) ², a saber: i) vertente profissional e ética; ii) desenvolvimento do ensino e da aprendizagem; iii) participação na escola e relação com a comunidade escolar e iv) desenvolvimento e formação profissional ao longo da vida. Numa das suas recomendações (Nº3/CCAP/2008), o Conselho Científico para a Avaliação de Professores (CCAP) (órgão nacional que monitoriza o processo de implementação da avaliação do desempenho docente e formula orientações e recomendações a este respeito) propõe a definição de padrões nacionais a partir do ano de 2009/2010 com base nas práticas desenvolvidas nas escolas, no debate e na investigação.

A avaliação de desempenho realiza-se de dois em dois anos, estando previsto o fim do primeiro ciclo de avaliação em Dezembro de 2009. O Decreto Regulamentar nº2/2008 especifica os procedimentos e o processo de avaliação de desempenho cujos traços principais se apresentam de seguida.

² O Decreto-Lei n.º 241/2001, de 30 de Agosto de 2001, aprova os perfis específicos de desempenho profissional dos educadores de infância e dos professores do 1.º ciclo do ensino básico e também constitui um referente para a avaliação.

TABLA 1. PROCEDIMENTOS E O PROCESSO DE AVALIAÇÃO DE DESEMPENHO, PORTUGAL

Pediodicidade	De dois em dois anos
Principais efeitos	Progressão na carreira
Elementos de referência para a avaliação	<p>* Objectivos e metas fixados no projecto educativo e no plano anual de actividades da escola ou agrupamento de escolas</p> <p>* Indicadores de medida previamente estabelecidos pelo agrupamento ou escola, nomeadamente quanto ao progresso dos resultados dos alunos e a redução das taxas de abandono escolar tendo em conta o contexto socioeducativo</p>
Avaliadores	<p>* Coordenador de departamento (pode delegar as suas competências de avaliador noutros professores titulares)</p> <p>* Presidente do Conselho Executivo/Director (pode delegar as suas competências de avaliador noutros membros da comissão directiva)</p> <p>No caso do coordenador de departamento, os avaliadores são o presidente do conselho executivo ou director e um inspector com formação científica na área do departamento do avaliado.</p> <p>O Director de escola ou Presidente do Conselho Executivo é avaliado pelo Director Regional de Educação. Os restantes membros do Conselho Executivo são avaliados pelo Director de Escola ou Presidente do Conselho Executivo.</p>
Aspectos a avaliar	<p>* O coordenador de departamento avalia o envolvimento e a qualidade científico-pedagógica do docente, tendo em conta os seguintes parâmetros: i) preparação e organização das actividades lectivas; ii) realização das actividades lectivas; iii) relação pedagógica com os alunos; iv) processo de avaliação das aprendizagens dos alunos. São observadas pelo menos três aulas (correspondentes a unidades didácticas diferentes) em cada ano escolar.</p> <p>* O órgão de direcção executiva avalia os seguintes aspectos: i) nível de assiduidade; ii) serviço distribuído; iii) progresso dos resultados escolares esperados para os alunos e redução das taxas de abandono escolar, tendo em conta o contexto socioeducativo; iv) participação dos docentes na escola ou agrupamento com base no número de actividades constantes do projecto curricular de turma e plano anual de actividades e qualidade e importância da intervenção do docente para o cumprimento dos objectivos; v) acções de formação contínua que incidam sobre conteúdos de natureza científico-didáctica relacionados com a matéria curricular que leccionam e com as necessidades do agrupamento ou da escola; vi) exercício de cargos ou funções de natureza pedagógica; vii) dinamização de projectos de investigação, desenvolvimento e inovação educativa.</p>
Procedimentos e instrumentos	<p>* A recolha de dados para efeitos de avaliação é feita através de instrumentos de registo normalizados, os quais são elaborados e aprovados pelo Conselho Pedagógico⁽¹⁾ dos agrupamentos ou escolas com base nas recomendações do Conselho Científico para a Avaliação de Professores (CCAP).</p> <p>* Fixação de objectivos individuais (por acordo entre avaliadores e avaliado; em caso de desacordo prevalece a posição dos avaliadores). Os objectivos individuais são formulados tendo em conta: i) a melhoria dos resultados escolares dos alunos; ii) a redução do abandono escolar; iii) a prestação de apoio à aprendizagem dos alunos, incluindo aqueles com dificuldades de aprendizagem; iv) a participação nas estruturas de orientação educativa e dos órgãos de gestão; v) a relação com a comunidade; vi) a formação contínua adequada ao cumprimento de um plano individual de desenvolvimento profissional do docente; vii) a participação e dinamização de projectos e actividades constantes no plano anual de actividades e de outros projectos e actividades extracurriculares.</p> <p>* O processo de avaliação inclui as seguintes fases: i) preenchimento da ficha de auto-avaliação; ii) preenchimento das fichas de avaliação pelos avaliadores; iii) conferência e validação das propostas de avaliação com menção qualitativa de Excelente, Muito Bom e Insuficiente, pela comissão de coordenação da avaliação; iv) realização da entrevista individual dos avaliadores com o respectivo avaliado; v) realização da reunião conjunta dos avaliadores para atribuição da avaliação final.</p>
Classificação final	<p>Excelente – de 9 a 10 valores Muito Bom – de 8 a 8,9 valores Bom – de 6,5 e 7,9 valores</p> <p>Regular – de 5 e 6,4 valores Insuficiente – de 1 e 4,9 valores</p> <p>Existe, no entanto, um sistema de quotas em cada escola para as menções de Excelente e Muito Bom (a serem fixadas em função da avaliação externa da escola)⁽²⁾.</p>
Coordenação do processo	<p>A comissão de coordenação da avaliação do desempenho criada em cada escola ou agrupamento integra o presidente do conselho pedagógico e quatro membros do mesmo conselho com a categoria de professor titular, designados pelo conselho pedagógico. Compete a esta comissão validar as menções finais de Excelente, Muito Bom e Insuficiente.</p> <p>Foi criado o Conselho Científico para a Avaliação de Professores que monitoriza o processo de implementação da avaliação do desempenho a nível nacional (Decreto Regulamentar 4/2008, de 5 de Fevereiro).</p>

⁽¹⁾ O Conselho Pedagógico é o órgão de coordenação e orientação educativa da escola, nomeadamente, nos domínios pedagógico-didáctico, de orientação e acompanhamento dos alunos e da formação inicial e contínua do pessoal docente e não docente. Neste órgão participam representantes das estruturas de orientação educativa e dos serviços de apoio educativo, das associações de pais e encarregados de educação, dos alunos no ensino secundário, do pessoal não docente e dos projectos de desenvolvimento educativo. O presidente do Conselho Executivo é também membro deste órgão.

⁽²⁾ As percentagens máximas para a atribuição de Excelente e Muito Bom em cada agrupamento de escolas ou escola são 5% e 20%, respectivamente. Contudo, estas percentagens podem ser mais elevadas, em função da avaliação externa da escola ou agrupamento. Por exemplo, no caso de cinco classificações de Muito Bom (classificação máxima) nos domínios e factores que fazem parte da avaliação externa das escolas (i) resultados; ii) prestação de serviço educativo; iii) organização e gestão escolar; iv) liderança e v) capacidade de auto-regulação e melhoria da escola) as percentagens para atribuição de Excelente e Muito Bom passam a ser, respectivamente, 10% e 25% (ver Despacho nº 20131/2008).

O Decreto-Lei 2/2008, de 10 de Janeiro, estipula ainda que os professores têm “direito à avaliação do seu desempenho a qual deve contribuir para o seu desenvolvimento profissional” (artigo 11º). Os docentes, também se declara no mesmo documento, têm direito a que lhes sejam “garantidos os meios e condições necessários ao seu desempenho, em harmonia com os objectivos” sendo um dever dos docentes “proceder à respectiva auto-avaliação como garantia do envolvimento activo e responsabilização no processo avaliativo e melhorar o seu desempenho em função da informação recolhida durante o processo de avaliação”. Por outro lado, é garantido aos docentes o conhecimento dos objectivos, fundamentos, conteúdo e funcionamento do sistema de avaliação do desempenho, a quem se reconhece o direito de reclamação e recurso. A auto-avaliação é obrigatória e visa “envolver o avaliado no processo de avaliação, de modo a identificar oportunidades de desenvolvimento profissional e de melhoria do grau de cumprimento dos objectivos fixados” (artigo 16º, Decreto-Lei 2/2008).

Nesta secção descrevemos, de forma necessariamente breve, a avaliação do desempenho no contexto português, ainda em fase de implementação. No ponto seguinte, damos conta de alguns aspectos do seu processo de implementação ao longo dos últimos dois anos, nomeadamente os sucessivos ajustes que foram sendo feitos à proposta acima descrita, através de dois processos de “simplificação” do modelo (em 2007/2008 e em 2008/2009), como forma de responder ao clima de contestação e turbulência em torno da sua implementação, em grande parte, associado ao aumento da burocracia e do volume de trabalho nas escolas e a dificuldades de gestão e concretização do modelo tendo em conta as condições de trabalho das escolas e dos professores.

3.3. O processo de implementação: tensões, contestação e resistência

A literatura neste domínio chama a atenção para a importância das percepções dos professores e para a complexidade da dimensão social na implementação de uma qualquer medida política, mais ainda numa matéria sensível e complexa como é o caso da avaliação de professores (Van der Vegt, Smyth, e Vandenberghe, 2001; Fullan, 2001; Flores, 2005; Tuytens e Devos, 2008). A investigação aponta para a necessidade de atender ao sentido que os actores envolvidos no processo de implementação atribuem a uma determinada medida, aos seus valores e emoções, bem como às interacções sociais e aos contextos em que a mudança vai ser implementada (Timperley e Robinson, 1997; van der Berg, Vandenberghe, e Sleegers, 1999; Spillane, Reiser, e Reimer, 2002). Numa recente revisão de alguns marcos referenciais para a avaliação do desempenho, Vaillant (2008) identifica os factores de natureza política, conceptual e operacional que podem facilitar ou obstaculizar este mesmo processo, chamando a atenção para, entre outros aspectos, a necessidade de ter em conta as variáveis contextuais na implementação de um determinado sistema de avaliação de desempenho, a adequação dos instrumentos de avaliação, a necessidade de os avaliadores serem reconhecidos e a importância do *feedback*. Também Tuytens e Devos (2008), num estudo recente realizado na Bélgica, identificaram a influência do director da escola nas percepções dos professores acerca do novo sistema de avaliação de professores, corroborando outros trabalhos empíricos neste domínio (Retallick e Fink, 2002; Kertsen e Israel, 2005). Neste sentido, para a eficácia e sucesso de um determinado sistema de avaliação, é necessário ter em consideração quer o conteúdo do sistema de avaliação, quer o contexto em que este vai ser implementado (Peterson e Comeaux, 1990).

Em Portugal, as medidas tomadas pelo Ministério da Educação, no que respeita à carreira e à avaliação dos professores, implicaram mudanças profundas com implicações não só em termos pessoais e profissionais para os professores, mas também no que se refere à escola enquanto organização,

sobretudo às relações de trabalho entre os professores, que passaram a basear-se nos princípios da hierarquização e diferenciação e não na paridade.

A alteração das regras de acesso e de progressão na carreira e o novo sistema de avaliação de desempenho têm gerado descontentamento e mal-estar nos professores tendo levado a uma forte contestação e resistência, os quais se traduziram em duas greves e duas grandes manifestações de protesto (organizadas pelos professores e pelos sindicatos e que contaram com mais de 100.000 docentes nas ruas de Lisboa) contra a diferenciação, a hierarquização e a avaliação. A existência de uma “carreira única” representava uma importante conquista para os sindicatos e constituía para os professores uma garantia de segurança, de justiça e de colegialidade, apesar de muitos professores desejarem a diferenciação. A burocracia, as reuniões, o aumento do volume de trabalho eram alguns dos aspectos mais negativos identificados pelos professores, a par da falta de reconhecimento e de formação adequada dos avaliadores.

De um modo geral, os últimos anos ficaram marcados por um ambiente de turbulência, agitação e contestação com implicações no trabalho das escolas e dos professores, o que ficou bem visível nos media, dando conta dos protestos e do clima de mal-estar e de ansiedade. Face a este cenário, e como resposta a esta situação – a de que não havia condições nas escolas para implementar o novo modelo de avaliação do desempenho docente –, o Ministério da Educação procedeu a duas “simplificações” no processo de avaliação. A primeira ocorreu na sequência de um Memorando de Entendimento celebrado com os sindicatos e introduziu uma versão simplificada para os professores a serem avaliados (nomeadamente os que necessitavam da atribuição da avaliação de desempenho para progredirem na carreira ou para renovação de contrato ou celebração de novo contrato, no ano lectivo de 2007/2008 (Decreto Regulamentar nº 11/2008, 23 de Maio)). Assim, esta versão simplificada previa os seguintes procedimentos: i) ficha de auto-avaliação; ii) avaliação por parte do órgão de direcção executiva em relação a: a) nível de assiduidade; b) cumprimento do serviço atribuído e c) acções de formação contínua. O segundo processo de simplificação ocorreu recentemente (Decreto Regulamentar nº 1-A/2009), depois de um conjunto de protestos, manifestações e duas greves por parte dos professores e dos sindicatos contra a aplicação do novo modelo de avaliação do desempenho docente. Foram identificados três problemas centrais: i) a existência de avaliadores de áreas disciplinares diferentes das dos avaliados; ii) a burocracia dos procedimentos existentes e iii) a sobrecarga de trabalho inerente ao processo de avaliação. Assim, o governo decidiu tomar um conjunto de medidas no sentido de simplificar o processo e ultrapassar estes problemas (a ser aplicado até ao final do primeiro ciclo de avaliação que termina em Dezembro de 2009), nomeadamente: i) garantir que os professores são avaliados por avaliadores da mesma área disciplinar; ii) dispensar os critérios dos resultados escolares e das taxas de abandono, no presente ano lectivo, tal como havia sido recomendado pelo Conselho Científico para a Avaliação de Professores (CCAP); iii) dispensar as reuniões entre avaliadores e avaliados sempre que haja acordo tácito; iv) a avaliação a cargo dos coordenadores de departamento só existe se os avaliados assim o requererem (incluindo observação de aulas), mas esta é uma condição necessária para a obtenção das classificações de Excelente e Muito Bom; v) reduzir para 2 em vez de 3 o número de aulas observadas, embora o avaliado possa requerer uma terceira aula observada; vi) dispensar da avaliação os professores que reúnam condições legais para se aposentarem até 2010/2011 ou requeiram a reforma antecipada; vii) dispensar da avaliação os docentes das áreas profissionais e vocacionais, a não ser que estes o desejem; viii) simplificar o regime de avaliação dos professores avaliadores e compensar a sua sobrecarga de trabalho.

De um modo geral, estas alterações e reajustamentos na aplicação do modelo da avaliação do desempenho resultaram, em grande medida, da falta de condições (incluindo tempo) para a sua operacionalização, a que não está alheio o facto de ter existido uma generalização do modelo sem adequada experimentação (e investigação), sobretudo num contexto marcado pela ausência de uma cultura de avaliação. Por outro lado, é de destacar a observação de aulas – um elemento chave num sistema de avaliação de professores – que deixou de ser obrigatória, nesta fase de transição, o que constitui um retrocesso numa das dimensões centrais da avaliação do desempenho. Outros aspectos críticos relacionam-se com a falta de reconhecimento e de formação adequada dos avaliadores e a burocracia que este modelo implica. Estas questões tornaram-se visíveis na contestação e resistência dos professores ao novo sistema de avaliação de desempenho, especialmente porque o viram como algo “imposto” e “contra os professores”. Tal pode relacionar-se com o modo como foi elaborado e implementado, quer em termos de *timing*, quer, como já se disse, em termos das condições (inexistentes) para a sua aplicação e generalização. Tal remete ainda para a ausência de discussão, experimentação, informação e formação adequadas numa perspectiva de mudança entendida mais como um facto ou como uma decisão e não como um processo que requer um sentido de apropriação e a consideração, não só do conteúdo, mas também dos contextos (e dos actores) em que a mudança se vai operar (Peterson e Comeaux, 1990). Por outro lado, trata-se de uma questão sensível e complexa que desafia as culturas profissionais dos professores (e os seus processos de socialização profissional) que requer a participação e o envolvimento dos docentes a quem a avaliação interessa em primeiro lugar. Este é, aliás, um dos aspectos-chave que tem sido reiterado na literatura neste domínio no sentido da eficácia e sucesso dos sistemas de avaliação. A este respeito, Assael e Pavez (2008:52) consideram que, mesmo no caso chileno, em que se registou uma construção participativa do sistema de avaliação de desempenho docente, a maioria dos professores encaram-no como uma imposição a que devem sujeitar-se “não incorporando as práticas avaliativas na sua própria cultura pedagógica como uma forma de reflexão sobre a sua prática e de desenvolvimento profissional”. Os mesmos autores identificam um conjunto de problemas no processo de implementação do sistema de avaliação do desempenho docente no Chile, cinco anos volvidos após a sua aprovação e implementação, destacando-se, entre outros, a desconfiança relativamente ao modo como os pares ou superiores hierárquicos avaliam; carga de trabalho associada à avaliação (que impede que esta se torne numa experiência formativa); ausência de planos de desenvolvimento profissional no sentido de ultrapassar e de melhorar os aspectos mais débeis identificados na avaliação; visão da avaliação como acto isolado, o que não concorre para a apropriação de uma cultura avaliativa de carácter formativo; etc.

Também em Portugal, num estudo recente sobre a auto-avaliação no contexto da avaliação do desempenho docente, realizado durante o processo de implementação do novo modelo de avaliação, Ribeiro (2008) concluiu que as expectativas dos professores face ao novo sistema de avaliação eram baixas, prevalecendo uma perspectiva negativa associada a factores de desigualdade e ao impacto negativo no que diz respeito às relações entre colegas, à competição e à burocracia. Eis alguns testemunhos: “Acho que haverá desigualdades, e suscita querelas entre colegas.”; “Penso que a avaliação de desempenho não vai trazer melhorias para o processo ensino-aprendizagem e que tem, apenas, como objectivo a penalização dos docentes, impedindo-os de progredir.”; “Muito competitiva, geradora de “poderes” (...) trazendo menos qualidade ao processo ensino/aprendizagem.”

Alguns professores assumiram uma posição céptica apontando para a ambiguidade e dúvida em termos dos efeitos do novo sistema de avaliação, demonstrando preocupação em relação aos propósitos e aos

processos de implementação da avaliação do desempenho e à sua repercussão na prática: "Com alguma apreensão. Espero que o fim essencial da avaliação do desempenho não seja a penalização dos docentes, mas o investimento na sua formação pessoal e profissional.", "A avaliação é absolutamente indispensável, mas a falta de regulamentação levanta-me sérias dúvidas quanto à sua eficácia." Apenas uma minoria manifestou uma visão positiva relativamente ao novo modelo de avaliação de desempenho docente, que surge associada à articulação entre propósitos formativos e sumativos, orientados para a prestação de contas, que os professores viam como um dos aspectos positivos do novo modelo. No entanto, de um modo geral, o perfil dos avaliadores, a natureza da comunicação entre avaliadores e avaliados e a necessidade de elaborar instrumentos adequados e fiáveis constituíam preocupações manifestadas pelos docentes que participaram neste estudo. O cepticismo em relação aos efeitos do novo sistema de avaliação e a falta de reconhecimento social da profissão docente aparecem ainda como dois aspectos críticos identificados pelos professores no processo de implementação do novo modelo de avaliação do desempenho docente (Ribeiro e Flores, 2008). O próprio Conselho Científico para a Avaliação de Professores (CCAP) reconheceu "as condições adversas em que o processo se iniciou" (Recomendações Nº2/CCAP/2008:2) e "a ruptura" que o modelo instituído implica ao nível das práticas de avaliação.

4. CONCLUSÃO

Globalmente, o processo de implementação do novo modelo de avaliação tem sido marcado pela resistência, contestação e turbulência, embora, de um modo geral, se assuma a necessidade de alterar o anterior modelo que se baseava na auto-avaliação. Fazendo um balanço do sistema de avaliação de desempenho docente no Chile, conseguido após vários anos de discussão e negociação, Avalos e Asseal (2006, p. 265) identificam um conjunto de recomendações com base nas lições aprendidas a partir do caso chileno: i) ampla participação dos vários *stakeholders* (agentes educativos), especialmente dos professores; ii) formulação participada de um sistema de critérios de avaliação centrado nas tarefas e competências docentes; iii) experimentação de uma variedade de modelos e de procedimentos; iv) articulação da avaliação de desempenho com outras políticas educativas, nomeadamente no que se refere às oportunidades para o desenvolvimento profissional de professores; v) resistir à tentação de "apressar" o design e implementação do processo de avaliação e estar aberto a mudanças e reajustamentos. Partindo destas questões e olhando para o caso português, podemos dizer que elas foram negligenciadas, sobretudo se tivermos em conta que se tratou de um modelo generalizado sem adequada (in)formação, experimentação e validação, não tendo, neste caso, a Administração Central "resistido à tentação", para utilizar a expressão das autoras, de implementar o sistema de avaliação com a maior celeridade possível. Assim, o grau de participação e de envolvimento não foi o adequado, não só pela natureza sensível e complexa da questão, mas tendo também em conta o facto de não existir uma cultura de avaliação. A literatura aponta para a importância da participação e do sentido de apropriação para o sucesso e eficácia dos sistemas de avaliação (ver, por exemplo, Avalos e Assael, 2006). De um modo geral, estamos perante um modelo de avaliação eminentemente burocrático e sumativo com efeitos na progressão na carreira docente. Entre as questões mais críticas, para além da burocracia, destacam-se o sistema de quotas (e concurso para professor titular numa fase transitória); a falta de reconhecimento e de formação dos avaliadores (entre o *design* e a implementação não houve tempo para uma formação relevante e adequada); as condições para a sua implementação, pelas dimensões que

envolve e pelas repercussões que implica; o sistema de *follow-up* e as oportunidades de desenvolvimento profissional, etc.

Desta breve reflexão podemos identificar um conjunto de riscos que a implementação deste modelo (e algumas das suas dimensões) encerra, nomeadamente ao nível das relações profissionais entre os docentes e dos seus efeitos. Em vez da colaboração entre os docentes (expressões reiteradas no enquadramento e justificação do novo modelo de avaliação), o risco é que se acentuem o individualismo e a competição, não se desafiando nem alterando as culturas profissionais dos professores (e processos de socialização). Por outro lado, existe o risco de este modelo não conduzir a mudanças significativas e não produzir efeitos reais ao nível do desenvolvimento profissional dos professores, da melhoria do ensino e da aprendizagem dos alunos e da melhoria da escola. Pelo contrário, há sinais de maior desmotivação, desgaste e desinvestimento (por exemplo, olhando ao número de aposentações, mesmo em casos de significativa penalização em termos de salário). Aliás, o próprio Conselho Científico para a Avaliação de Professores (CCAP) alertava, numa das suas recomendações (Nº2/CCAP/2008:1), para o “risco da avaliação se constituir num acto irrelevante para o desenvolvimento profissional de professores, sem impacto na melhoria das aprendizagens dos alunos” resultante da “burocratização excessiva, da emergência ou reforço de conflitualidades desnecessárias e do desvio das finalidades formativas e reguladoras que um processo de avaliação de desempenho profissional deve conter”.

Torna-se necessário, portanto, clarificar os propósitos da avaliação de desempenho e os seus efeitos reais se se pretende melhorar o ensino, os professores e as escolas. E é sobretudo fundamental a existência de um sentido de apropriação e um clima de confiança e de credibilidade para que a sua implementação seja eficaz.

Resta aguardar e ver quais serão os efeitos deste modelo de avaliação no futuro. Contudo, os indicadores apontam para situações de algum mal-estar e tensão em muitas escolas. Questões ligadas à natureza da comunicação (um elemento central para que um sistema de avaliação resulte), a qualidade do *feedback*, a relação entre avaliadores e avaliados, o tempo necessário para conduzir o processo de avaliação (incluindo observação de aulas), a existência de oportunidades relevantes de desenvolvimento profissional, etc. são algumas das áreas críticas que merecem mais reflexão e investigação.

Numa lógica de avaliação enquanto estratégia de desenvolvimento e de melhoria, a participação dos docentes é fundamental e decisiva para a compreensão das suas práticas, para a reflexão sobre o processo de ensino-aprendizagem, para a promoção da sua própria aprendizagem e desenvolvimento profissional, o que pode (e deve) incluir a participação, ajuda e colaboração de colegas a vários níveis e em diversos âmbitos à luz do lema “you don’t need to be ill, to get better” (não é preciso estar doente para melhorar), ou seja, num esforço constante de promoção da qualidade do ensino e da valorização da aprendizagem a partir da avaliação.

REFERÊNCIAS BIBLIOGRÁFICAS

- Asseal, J. e Pavez, J (2008) La construccion e implementacion del sistema de evaluacion del desempeño docente chileno: principales tensiones y desafios. *Revista Iberoamericana de Evaluacion Educativa*, 1 (2), 41-55.
- Avalos, B. (2004) Teacher regulatory forces and accountability policies in Chile: from public servants to accountable professionals. *Research Papers in Education*, 19(1), pp. 67-85.

- Avalos, B. e Assael, J. (2006) Moving from resistance to agreement: The case of the Chilean teacher performance evaluation. *International Journal of Educational Research*, 45, pp. 254-266.
- Chow, A.P.Y., Wong, E.K.P., Yeung, A.S. e Mo, K.W. (2002) Teachers' perceptions of appraiser-appraisee relationships, *Journal of Personnel Evaluation in Education*, 16 (2), 85-101.
- Darling-Hammond, L. (2008) Recognizing and Enhancing Teacher Effectiveness: A Policymaker's Guide, in *Council of Chief State School Officers. Strengthening teacher quality in high-need schools- policy and practice. Washington, DC.*
- Darling-Hammond, L.; Wise, A. E. e Pease, S. R. (1983) Teacher Evaluation in the Organizational Context: A Review of the Literature. *Review of Educational Research*, 53(3), 285-328
- Day, C. (1992) Avaliação do desenvolvimento profissional dos professores. In A. Estrela e A. Nóvoa (Orgs.) *Avaliações em Educação: novas perspectivas*. Lisboa: Educa, pp. 89-104
- Day, C. (1999) *Developing Teachers. The Challenges of Lifelong Learning*. London: Falmer Press.
- Day, C. (2004) *A Paixão do ensino*. Porto: Porto Editora.
- Day, C.; Flores, M. A. e Viana, I. (2007) Effects of national policies on teachers' sense of professionalism: findings from an empirical study in Portugal and in England. *European Journal of Teacher Education*, 30 (3), 249-266.
- Decreto Regulamentar 4/2008, de 5 de Fevereiro
- Decreto Regulamentar nº 11/2008, 23 de Maio
- Decreto Regulamentar nº 1-A/2009
- Decreto Regulamentar nº 2/2008, de 10 de Janeiro
- Decreto-Lei 2/2008, de 10 de Janeiro
- Decreto-Lei 240/2001, de 30 de Agosto
- Decreto-Lei n.º 241/2001, de 30 de Agosto
- Decreto-Lei nº 139-A/90, 28 de Abri
- Decreto-Lei nº 15/2007, de 19 de Janeiro
- Despacho nº 20131/200
- Flores, M. A. (2003) Teacher Professionalism in Portugal: Theory and Practice, in G.G. Markea and G. T. Papanikos (Eds) *Global Issues of Education*, Proceedings from the 5th International Conference on Education, Athens, Greece, May 23-24 (Vol 2), pp.235-242.
- Flores, M. A. (2005) Teachers' views on recent curriculum changes: tensions and challenges. *The Curriculum Journal*, 16 (3), 401-413.
- Flores, M. A. e Viana, I. C. (2004) The Impact of Recent Elementary School Curriculum Reorganisation on Teachers' Attitudes and Practices of Teaching, Comunicação apresentada na European Conference on Educational Research (ECER), Universidade de Crete, Grécia, Setembro 22-25.
- Flores, M. A., Day, C. e Viana, I. C. (2007) Profissionalismo Docente em transição. As identidades dos professores em tempos de mudança. Um estudo com professores ingleses e portugueses. In M. A. Flores e I. C. Viana (Orgs.) *Profissionalismo docente em transição. As identidades dos professores em*

- tempos de mudança*. Braga. CIEEd, pp. 7-45
- Fullan, M. (2001) *The new meaning of educational change*. (3rd edition) London. RoutledgeFalmer.
- Hadji, C. (1994a) A avaliação dos professores. Linhas directivas para uma metodologia pertinente, in A. Estrela e P. Rodrigues (org) *Para uma fundamentação da Avaliação em Educação*, Lisboa: Colibri, pp. 37-50
- Hadji, C. (1994b) *Avaliação: as regras do jogo*. Porto: Porto Editora
- Hargreaves, A. e Goodson, I. (1996) Teachers' Professional Lives: Aspirations and Actualities, in I. Goodson and A. Hargreaves (Eds) *Teachers' Professional Lives* London: Falmer Press, pp. 1-27.
- Hargreaves, Andy (1998). *Os professores em tempos de mudança. O trabalho e a cultura dos professores na idade pós-moderna*. Lisboa: MacGraw-Hill.
- Kertsen, T. A. & Israel, M. S. (2005) Teacher evaluation: principals' insights and suggestions for improvement. *Planning and Changing*, 36(1,2), 47-67.
- Lima, L. (2006) 'Administração da Educação e Autonomia das escolas', in *A Educação em Portugal (1986-2006). Alguns contributos de investigação*. Sociedade Portuguesa de Ciências da Educação, 5-54.
- Middlewood, D. e Cardno, C. (eds) (2001) *Managing Teacher Appraisal and Performance. A comparative Approach*. London, RoutledgeFalmer.
- Morgado, J. C. (2003) *Processos e Práticas de (Re)construção da autonomia curricular*. Tese de Doutoramento. Universidade do Minho.
- National Board for Professional Teaching Standards (1999) *What teachers should know and be able to do*. Washington, DC: NBPTS
- Nóvoa, A. (1992) Os professores e as histórias da sua vida. In A. Nóvoa (org.) *Vidas de Professores*. Porto: Porto Editora, pp. 11-30
- Pacheco, J. A. e Flores, M. A. (1999). *Formação e Avaliação de Professores* Porto: Porto Editora
- Palmer, J. P. (1998) *The Courage to Teach. Exploring the inner landscape of a teacher's life*. San Francisco, CA: Jossey-bass Inc, Publisher
- Peterson, P. e Comeaux, M. A. (1990) Evaluating the systems. Teachers' perspectives on Teacher Evaluation. *Educational Evaluation and Policy Analysis*, 12 (1), pp. 3-24.
- Princípios Orientadores para a definição dos padrões relativos às menções qualitativas. Recomendações N°3/CCAP/2008, Conselho Científico para a Avaliação de Professores.
- Princípios Orientadores sobre a organização do processo de avaliação do desempenho docente. Recomendações N°2/CCAP/2008. Conselho Científico para a Avaliação de Professores
- Retallick, J. & Fink, D. (2002) Framing leadership: contributions and impediments to educational change. *International Journal of Leadership in Education*, 5(2), 91-104
- Ribeiro, C. (2008) *Avaliação de desempenho no contexto educativo Português: Perspectivas dos professores sobre Auto-Avaliação*. Master Dissertation, University of Minho, Portugal.
- Ribeiro, C. e Flores, M. A. (2008) A auto-avaliação no contexto da avaliação do desempenho docente in Alves, M. P., Machado, E. A., Fernandes, J. A. & Correia, S. (orgs.) (2008). *Actas do Seminário de Avaliação: Contributos da investigação para as práticas quotidianas de desenvolvimento curricular das*

- escolas*. (Braga, CIEd, Universidade do Minho) ISBN 978-972-8746-60-5, pp. 506-534.
- Sachs, J. (2003) Teacher Professional Standards: controlling or developing teaching. *Teachers and Teaching: theory into practice*, 9 (2), pp. 175-186
- Silva, I. M. L. P. & Conboy, J. E. (2004). Representações dos professores sobre a avaliação do desempenho docente: o que avaliar? Como avaliar? *Revista Portuguesa de Educação*, XXIV, pp. 119-150
- Simões, G. (2000). *A Avaliação do Desempenho Docente, Contributos para uma análise crítica*. Lisboa,: Texto Editora.
- Smith, K. (2005) New Methods and Perspectives on Teacher Evaluation. Who evaluates what and for which purposes? In D. Beijaard, P. C. Meijer, G. Morine-Dersheimer and H. Tillema (Eds) *Teacher Professional Development in Changing Conditions*, Dordrecht: Springer, pp. 95-11
- Spillane, J. P.; Reiser, B.J. & Reimer, T. (2002) Policy implementation and Cognition: Reframing and refocusing implementation research. *Review of Educational Research*, 72 (3), 387-431.
- Stronge, J. H. (2006) *Evaluating teaching: a guide to current thinking and best practice*. (2nd edition) Thousand Oaks, CA: Corwin Press.
- Timperley, H.S. & Robinson, V. M. J. (1997) The problem of policy implementation: the case of performance appraisal . *School Leadership and Management*, 17 (3), 333-345.
- Tuytens, M. & Devos, G. (2008) New Educational policy on teacher evaluation: the influences of school leadership on Policy perception of teachers. Comunicação apresentada na Annual Meeting of the American Educational Research Association (AERA), New York, March 25-27 2008.
- Vaillant, D. (2008) Algunos marcos referenciales en el evaluacion del desempeño docente. *Revista Iberoamericana de Evaluacion Educativa*, 1 (2), 7-22.
- Van der Berg, R., Vandenberghe, R. & Sleegers, P. (1999) Management of innovations from a cultural-individual perspective. *School Effectiveness and School Improvement*, 10, 321-351.
- Van der Vegt. R., Smyth, L.F. & Vandenberghe, R. (2001). Implementing educational policy at the school level. Organization, dynamics and concerns. *Journal of Educational Administration*, 39, 8-23.
- Veiga Simão, A. M., Caetano, A. P., O'Meara, J. & Flores, M. A. (2003) Toward a conceptualisation of Teacher Change: contexts, processes and factors, Paper presented at the 48th World Assembly of the International Council on Education for Teaching (ICET), Melbourne, Australia, 3-7 July 2003.
- Whitty, G. (2000) Teacher Professionalism in New Times, *Journal of In-Service Education*, 26 (2) 281-295.

¿CÓMO EVALUAR LA CONVIVENCIA ESCOLAR? LA MIRADA DESDE LOS ESTUDIANTES

Álvaro Carrasco

Revista Iberoamericana de Evaluación Educativa 2009 - Volumen 2, Número 1

<http://www.rinace.net/riee/numeros/vol2-num1/art14.pdf>

Fecha de recepción: 2 de marzo de 2009

Fecha de comunicación de dictamen: 11 de abril de 2009

Fecha de aceptación: 22 de abril de 2009

Generalmente, las evaluaciones de logro han delimitado el concepto de calidad a los aprendizajes de las materias del currículo, como las matemáticas y el lenguaje. Es posible argumentar que esta reducción del concepto de aprendizaje es necesaria porque estos son los conocimientos básicos que se espera que la escuela entregue y en los que tendría mayor injerencia; porque es difícil evaluar ciertos aspectos educativos, como la formación en valores, con instrumentos estandarizados; porque, se asume que el éxito futuro está relacionado fundamentalmente con las destrezas cognitivas de orden superior. Ciertamente hay validez en dichos argumentos, sin embargo también es cierto que igualmente básico es aprender a convivir, que las escuelas son, muchas veces, una burbuja para los jóvenes entre la violencia y el riesgo de sus hogares y comunidades, y que, desde la teoría y la investigación (ver por ejemplo: Cassassus y otros, 2000), se destacan la importancia de la inteligencia emocional o el clima de la escuela como predictores de éxito académico.

Siempre ha estado en el foco de la educación la formación “integral” del niño. Por integralidad, además de los aprendizajes cognitivos más obvios, se han entendido diversos aspectos: educación en valores, educación del carácter, educación cívica, educación para la paz, destrezas para la vida, destrezas cognitivas, derechos humanos, inteligencia emocional, prosocialidad, resolución de conflictos, educación cívica.

En buena medida esta multiplicidad es acertada en tanto que refiere a fenómenos interrelacionados e interdependientes, que forman un dinamismo complejo en el individuo y sus vínculos. Sin bien esta mixtura conceptual no es, necesariamente, un impedimento muy importante desde el punto de vista pedagógico, sí presenta un obstáculo cuando se trata de evaluar la “integralidad” de la educación. De allí la relevancia de acotar y definir el fenómeno.

En los últimos años, ha tomado fuerza el concepto de convivencia escolar que ofrece límites suficientemente claros, a la vez que adecuada amplitud, para permitir un buen acercamiento a la diversidad de habilidades y relaciones implicadas. La formulación de la Política de Convivencia en Chile abarca un ámbito de la experiencia escolar que se refuerza y complementa con las definiciones en otros programas y políticas. En la primer parte de este trabajo se identifican y revisan, brevemente, dichas políticas.

La propuesta de evaluar la convivencia descansa en la convicción de la necesidad de complementar y mejorar las evaluaciones de la calidad educativa. Es importante evaluar la convivencia, en tanto que si no se lo hace, si se evalúa y se definen consecuencias, exclusivamente, para los resultados de evaluaciones académicas, los estímulos, esfuerzos y recursos estarán orientados de forma desequilibrada. La segunda parte de este trabajo esboza una propuesta sobre qué evaluar en el ámbito de la convivencia escolar. Específicamente se ofrece un instrumento para que los estudiantes de educación secundaria expresen su opinión sobre la convivencia en sus establecimientos.

1. POLÍTICAS DE CONVIVENCIA ESCOLAR

Existe un conjunto de marcos legales que legitiman y obligan la atención a la convivencia en la escuela: la Constitución Política de la República de Chile; la Declaración Universal de los Derechos Humanos; la Convención sobre los Derechos del Niño; la Ley Orgánica Constitucional de Enseñanza; los Decretos Supremos de Educación, en lo que respecta a los Objetivos Fundamentales Transversales; el Instructivo

Presidencial sobre Participación Ciudadana, la Política de Participación de Padres, Madres y Apoderados en el sistema educativo y las normativas para el funcionamiento de Consejos Escolares.

El conjunto de políticas propone principios para un cambio en la forma tradicional de concebir la convivencia escolar y a los miembros de la comunidad escolar: más participación en un estilo de convivencia democrático; más comprensión de la necesidad de vivir entre diversos, todos iguales en dignidad, respetuosos de las diferencias; más atención al desarrollo personal; desarrollo de la capacidad reflexiva, crítica y ética (Valoras, 2006).

A continuación se presentan brevemente algunas de las políticas más relevantes a tener en cuenta en una evaluación de la convivencia.

1.1. Política de participación de los apoderados

El Ministerio de Educación de Chile plantea como objetivo de su Política de Participación de Padres, Madres y Apoderados (2002a): generar condiciones institucionales y jurídicas que garanticen la promoción, fomento y desarrollo de la participación amplia, diversa y organizada de padres, madres y apoderados/as y de la comunidad en el sistema educativo. Esta participación se realiza en alianza con los demás actores de la comunidad educativa, con el fin de contribuir en la construcción, desarrollo y cumplimiento de su proyecto institucional, en el marco de mejorar la calidad de la educación.

Desde dicho objetivo general la Política define 4 áreas de impulso estratégico y varias líneas de acción:

- Fomentar el Diálogo y la Alianza entre Actores.
- Formar y Apoyar a los Actores para la Participación en el Sistema Educativo.
- Habilitar Jurídica e Institucionalmente la Participación.
- Sensibilizar, Motivar y Difundir la Política Líneas de Acción.

1.2. Objetivos Fundamentales Transversales (OFT)

La reforma de la educación chilena ha incorporado, en el marco de un currículo flexible y descentralizado, la promoción de ciertos valores y actitudes asociadas con las formas de relaciones en el trabajo moderno, con la ciudadanía y la democracia. Con esto se pretende que la educación fortalezca, no solamente la dimensión cognitiva de los estudiantes sino, también, sus aspectos emotivo y social; en otras palabras, expresa la tradicional búsqueda de la "educación integral". Desde el diseño curricular se proponen algunos objetivos educativos fundamentales que todo colegio, al construir sus programas locales, debe incluir.

Los OFT se han estructurado en cuatro ámbitos que están estrechamente vinculados entre sí (Assael y otros, 2003:52-54):

- Crecimiento y Autoafirmación Personal.
- Desarrollo del pensamiento.
- Formación ética.
- El ámbito de la persona y su entorno.

Los Objetivos Fundamentales Transversales establecidos en el Marco Curricular, son uno de los fundamentos de la Política de Convivencia Escolar, de modo que se encuentran en una relación directa. Así, el logro de estos objetivos implica el logro de los propósitos de la Política de Convivencia y viceversa.

1.3. Política de convivencia

La Política de Convivencia Escolar (MINEDUC, 2002b), promueve la formación de valores para la convivencia y participación en la comunidad, a través del señalamiento de la obligación de formar en valores como el respeto por la diversidad, participación activa en la comunidad, colaboración, autonomía, solidaridad; erradicando las prácticas discriminatorias y abusivas.

La convivencia escolar es una construcción colectiva y dinámica entre los diferentes miembros de un establecimiento educacional, que tiene incidencia significativa en el desarrollo ético, socioafectivo e intelectual de alumnos y alumnas. Esta concepción no se limita a la relación entre las personas, sino que incluye las formas de interacción de los diferentes estamentos que conforman una comunidad educativa, por lo que constituye una construcción colectiva y es responsabilidad de todos los miembros y actores educativos sin excepción. Así mismo, varía de una comunidad escolar a otra de acuerdo a las características particulares de sus miembros y de las interrelaciones que entre ellos establecen. Por tanto, no admite un modelo único (MINEDUC, 2002b). No obstante, el Ministerio de Educación estima necesario explicitar un mínimo común para todas las comunidades educativas del país. Este mínimo es el que se quiere reflejar en los principios rectores de la Política de Convivencia Escolar. Los principios rectores, en su conjunto, sintetizan aspectos relevantes que se debieran cautelar en la cotidianeidad de todo establecimiento educacional:

Todos los actores de la comunidad educativa son sujetos de derecho.

- Los niños, niñas y jóvenes son sujetos de derecho.
- La educación como pleno desarrollo de la persona.
- Convivencia democrática y construcción de ciudadanía en la institución escolar.
- La convivencia escolar: un ámbito de consistencia ética.
- Respeto y protección a la vida privada y pública y a la honra de la persona y de su familia.
- Igualdad de oportunidades para niños y niñas, mujeres y hombres.
- Las normas de convivencia: un encuentro entre el derecho y la ética.

2. PROPUESTA DE EVALUACIÓN DE LA CONVIVENCIA ESCOLAR EN CHILE

¿Qué evaluar? Se podría tomar el camino de evaluar los resultados, es decir a los estudiantes en sus atributos relacionados con la convivencia escolar. Este camino presenta el problema de definir el constructo a medir; en algunas investigaciones y experiencias (por ejemplo en: Murillo y otros, 2007) se ha usado el concepto de autoestima, más específicamente, a veces, autoestima escolar; hay ocasiones¹ en que se ha optado por la inteligencia emocional o actitudes prosociales. En el caso de la autoestima, el concepto, aunque parcialmente útil, es muy indirecto y reducido para la amplitud de dimensiones que, como arriba se ha expuesto, pretende abarcar las acciones pro convivencia. Inteligencia emocional presenta la dificultad de su medición; una vía común son los autoreportes de conducta. Estos tienen la desventaja de que las respuestas pueden ser falseadas por los estudiantes en un intento de generar una

¹ Experiencia de evaluación de un grupo de colegios religiosos de Santiago en la que participó el autor de este artículo.

impresión positiva. El mismo problema enfrenta la medición de actitudes, además de la distancia entre la intención y la acción (Carrasco y otros, 1995). Hay algunos instrumentos desarrollados para medir habilidad que demuestren distintos niveles de inteligencia emocional esto mediante la presentación de algunas situaciones problemas que resolver². Esta parece ser una mejor estrategia, que no ha sido suficientemente explorada en Chile y que ameritaría mayor investigación y recursos. Sin embargo, un problema común a todos estos intentos es que se centran en los individuos. A diferencia de los aprendizajes académicos del currículo formal, en general, los jóvenes aprenden a convivir primero en sus familias, luego en sus barrios y también en la escuela. Si se sabe que, aun en áreas como matemáticas y lenguaje el mayor determinante de éxito es el capital familiar, sería excesivamente optimista atribuir a la escuela los resultados en la medición de las variables individuales pro-convivencia.

Por otra parte, la convivencia, tal como se ha definido más arriba, es una “construcción colectiva y dinámica” del grupo de individuos y sectores que conforman una comunidad escolar, por lo que es más apropiado que la evaluación atienda a los vínculos y perspectivas de los distintos actores. De allí entonces que esfuerzos por evaluar la convivencia deben incluir las diversas opiniones sobre las relaciones recíprocas y las acciones educativas claves en este ámbito. Idealmente la información proveniente de estudiantes, profesores y apoderados debería ser comparada de forma que pudieran identificarse las áreas de acuerdo y desacuerdo. Ambas áreas son importantes pero, desde el punto de vista de las escuelas, especial atención debería ponerse a las áreas de desacuerdo.

Con base a la experiencia, en esta propuesta se piensa la evaluación a partir de instrumentos de opinión, tipo cuestionario, pero en el caso de la exploración de las áreas de desacuerdo resulta más adecuadas técnicas cualitativas, como los grupos de discusión o entrevistas grupales.

Dimensiones a evaluar. De la revisión de las políticas chilenas relacionadas con convivencia escolar se desprenden las dimensiones del constructo a evaluar:

- participación de los padres;
- participación de los alumnos;
- objetivos fundamentales transversales;
- normas de convivencia;
- relaciones entre profesores, profesores-alumnos, profesores-apoderados;
- clima de aula;
- resolución de conflictos;
- violencia y matonaje (bullying).

Ciertamente cada una de estas dimensiones tiene profundidades y matices que son necesarios explorar antes de construir instrumentos para su evaluación. Dada la naturaleza de este trabajo refiero a los lectores, como una primera aproximación, al apartado de políticas sobre convivencia³.

² Ver por ejemplo: Mayer-Salovey-Caruso Emotional Intelligence Test (MSCEIT)

³ Varios de los textos sobre políticas de convivencias, y otros relacionados, han sido recopilados en la siguiente dirección electrónica:
http://cgjung.cl/component/docman/cat_view/19-convivencia-escolar

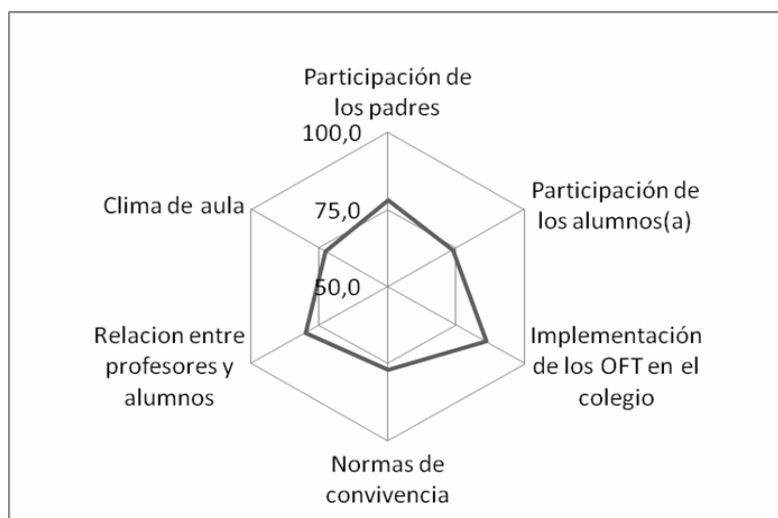
3. PROPUESTA DE UN INSTRUMENTO SOBRE CONVIVENCIA SEGÚN LOS ESTUDIANTES

3.1. Desarrollo del instrumento

Este instrumento se hizo teniendo en cuenta una revisión de las políticas chilenas sobre convivencia, fue elaborado en el marco de un proyecto de apoyo en convivencia a liceos prioritarios de la Región Metropolitana de Chile. A modo de prueba piloto, se realizó una revisión del instrumento por los estudiantes del curso Métodos Cuantitativos de la carrera de Psicología de la Universidad Alberto Hurtado. Este mismo grupo de estudiantes, bajo la supervisión del autor de este artículo, realizó (2008) una aplicación del cuestionario en todos los estudiantes de enseñanza media de un colegio particular subvencionado (no prioritario) de la comuna de Conchalí en Santiago (N=219).

Los resultados fueron presentados a las autoridades, profesores y apoderados del colegio quienes apreciaron la información entregada. Algunos años antes, se aplicó, en este mismo establecimiento, un instrumento de autoreporte de competencias socio-emocionales. En esa ocasión se compararon los datos de este colegio con los de otros establecimientos de Santiago ubicándoles en las zonas de altos puntajes promedios de la distribución. La convivencia escolar reportada, más recientemente, por los alumnos también resulta alta, habiendo una consistencia entre ambas aplicaciones.

EJEMPLO. PORCENTAJE POR SUB-ESCALA DE RESPUESTAS FAVORABLES SOBRE LA CONVIVENCIA ESCOLAR



Este gráfico resume las respuestas favorables (opciones de respuesta: "De acuerdo" y "Muy de Acuerdo") en afirmaciones relacionadas con las dimensiones exploradas en algunas sub-escalas del Cuestionario. Así por ejemplo, se informa que el 86% de los alumnos afirma participar de actividades relacionadas con los OFT del currículo. Este dato resulta consistente con la prioridad que acciones relacionadas con los OFT tienen en el proyecto educativo de este establecimiento. Sin embargo llamó la atención que un 28% de los alumnos estuviera en desacuerdo con las afirmaciones indicativas de un buen clima de aula.

3.2. Cuestionario para evaluar la convivencia escolar según los estudiantes

Se ofrece el instrumento como una versión preliminar, que puede resultar útil e informativa y que puede ser modificada y mejorada según estimen conveniente. Dicho instrumento se entrega en el Anexo del texto.

Este instrumento fue diseñado para ser aplicado a estudiantes de enseñanza media. El cuestionario está compuesto de varias sub-escalas:

- Información del alumno: preguntas 1 y 2.
- Involucramiento de los padres: preguntas 3.1 a 3.6.
- Participación de los estudiantes: 3.7 a 3.14.
- Actividades que implican los Objetivos Fundamentales Transversales: pregunta 4.
- Objetivos Fundamentales Transversales en el currículo formal y oculto: pregunta 5.
- Normas de convivencia: preguntas 6.1 a 6.7.
- Relaciones con profesores: preguntas: 6.8 a 6.13.
- Clima del aula: preguntas 6.14 a 6.18
- Conductas inadecuadas de parte de los profesores hacia los estudiantes: 7.1
- Conductas inadecuadas de parte de los estudiantes hacia los profesores: 7.2
- Resolución de conflictos: pregunta 8.
- Violencia, matonaje (*bullying*): pregunta 9.
- Clima familiar⁴: pregunta 10.

El análisis debe hacerse por sub-escala y además se pueden agrupar, en un índice total de convivencia, todas las sub-escalas que tienen un mismo tipo de respuesta (muy en desacuerdo...muy de acuerdo). En el sitio <http://cgjung.cl/cuestionario-convivencia> se ofrece la estructura de la base de datos en SPSS para digitar la información que se recoja, así como la sintaxis para realizar los análisis de los resultados.

Un problema detectado con esta versión guarda relación con su extensión⁵. El instrumento trata de explorar múltiples dimensiones de la convivencia, de allí su extensión. A pesar de que es un poco largo puede ser útil si solo se aplica este instrumento y se hace conciencia a los estudiantes sobre la importancia de responder con seriedad y sinceridad.

4. CONSIDERACIONES FINALES

Los principales indicadores de calidad educativa a nivel escolar han estado vinculados a los logros de las escuelas en pruebas de conocimientos estandarizadas. Otros aprendizajes trascendentes han sido menos atendidos produciéndose un desequilibrio perjudicial para todo el proceso educativo.

En Chile, considerando varias políticas educativas, es posible construir un marco conceptual coherente y amplio para evaluar la convivencia escolar. En este sentido resulta indispensable atender a la Política de Participación de Apoderados, al Currículo Nacional en sus Objetivos Fundamentales Transversales y a la Política de Convivencia Escolar.

⁴ Aunque esta no es una pregunta sobre convivencia escolar puede resultar útil a quienes les interese contratar con la situación del hogar de los estudiantes.

⁵ Recientemente, se aplicó una versión reducida de este instrumento, como una sub-escala de un instrumento más amplio, a una muestra representativa de estudiantes de liceos municipales de una comuna de la V región, los datos están siendo procesados.

Se puede evaluar a los individuos en relación a sus habilidades de convivencia. Un concepto útil en este sentido puede ser el constructo de inteligencia emocional. Los autoreportes son una buena estrategia pero presentan dificultades que serían superadas con tests de desempeño. El desempeño o medición de competencias de inteligencia emocional es más difícil de evaluar, aunque existen algunos instrumentos en el mundo anglosajón que sería interesante estudiar para su desarrollo en los países de habla hispana.

Sin embargo, por ser la convivencia un fenómeno interpersonal y eminentemente social resulta quizás más conveniente una evaluación que sea capaz de recoger y contrastar las percepciones de los distintos actores escolares. En este trabajo se presenta un cuestionario para estudiantes de secundaria que idealmente debería ser complementado con información del mismo tipo proveniente de profesores y apoderados. La superposición de estos datos puede ayudar a generar un panorama de la situación.

Ha de tenerse en cuenta que el objetivo fundamental de este tipo de evaluación no ha de ser el establecer comparaciones entre colegios. El propósito principal debería ser el generar información que permita a la escuela propiciar una reflexión participativa orientada a la acción positiva.

Para el autor de este trabajo resultó interesante constatar recientemente, en una serie de entrevistas en colegios municipales, que una de las variables más apreciadas por los apoderados es la buena convivencia escolar. Información relativa a estos procesos puede ayudar a los colegios a difundir logros atractivos que ahora son invisibles o que se comunican de forma asistemática.

REFERENCIAS BIBLIOGRÁFICAS

- Carrasco, Á.; Cárdenas, M.; Cornejo, R. & Galdámez, M. (1995). *Evaluación de la técnica del juego aplicada a una educación en derechos humanos que enfatice la promoción de actitudes y comportamientos prosociales*. Tesis para optar al Grado de Licenciados en Psicología. UCA, El Salvador.
- Casassus, J.; Cusato, S.; Froemel, J. E. F y Palafox, J. C. (2000). *Primer estudio internacional comparativo sobre lenguaje, matemática y factores asociados, para alumnos del tercer y cuarto grado de la educación básica (segundo informe)*. Laboratorio Latinoamericano de Evaluación de la Calidad de la Educación, UNESCO
- Assael, J.; Egaña, L.; Magendzo, A.; Santa Cruz, E.; Varas, R. (2003). *Reforma Educativa y Objetivos Fundamentales Transversales: Los Dilemas de la Innovación*. PIIE, Chile.
- MINEDUC. (2002a). *Política de participación de padres, madres y apoderados en el sistema educativo*. Obtenido en febrero del 2009 desde: <http://www.mineduc.cl/biblio/>
- (2002b). *Política de Convivencia Escolar*. Obtenido en febrero del 2009 desde: <http://www.mineduc.cl/biblio/>
- Murillo, F.J. (Coord.), Castañeda, E., Cueto, S., Donoso, J.M., Fabara, E., Hernández, M.L., Herrera, M., Murillo, O., Román, M. y Torres, P. (2007). *Investigación Iberoamericana sobre Eficacia Escolar*. Bogotá: Convenio Andrés Bello.
- Valoras (M. Isidora Mena, Neva Milicic, Claudia Romagnoli, Ana María Valdés). (2006). *Propuesta Valoras UC Potenciación de la política pública de Convivencia Social Escolar*. Obtenido en febrero del 2009 en: http://www.valorasuc.cl/documentos_valoras.htm

CUESTIONARIO

Este instrumento se comparte bajo una [Licencia Creative Commons de Atribución-No Comercial 2.0 Chile](#) 

En este cuestionario, nuestro interés consiste en obtener información sobre algunas características tuyas y de tu Colegio. Esta información es especialmente valiosa, por lo que solicitamos tu colaboración respondiendo a las preguntas que se te plantean. **No existen respuestas correctas ni incorrectas**; además, la información que nos proporciones es anónima, por lo que te pedimos que respondas lo más sinceramente posible.

INSTRUCCIONES

Por favor, lee atentamente y responde **todas** las preguntas. Recuerde que la información solicitada no pretende evaluarte, sino conocer mejor a tu Colegio.

La mayor parte de las preguntas están formuladas para que señales con una "X" en el cuadro o cuadros que correspondan. En otros casos, deberás escribir la respuesta en el espacio determinado para tales fines.

En muchas de las afirmaciones aparece la palabra padres; si tú no vives con tus dos padres o si vives con otras personas o familiares, debes entender por "padres", la o las personas que son responsables de ti ante tu Colegio.

De antemano te agradecemos tu colaboración

1. ¿Cuál es tu género?

Masculino ☐ Femenino ☐

2. ¿Quiénes viven en tu casa? Marca todas las alternativas que correspondan.

Madre	<input type="checkbox"/>
Madrastra	<input type="checkbox"/>
Padre	<input type="checkbox"/>
Padrastro	<input type="checkbox"/>
Hermanos o hermanas	<input type="checkbox"/>
Uno o más de los abuelos	<input type="checkbox"/>
Otros parientes (tíos, primos, etc.)	<input type="checkbox"/>
Otras personas que no sean parientes	<input type="checkbox"/>

3. ¿Cuán de acuerdo estás con las siguientes afirmaciones?

	Muy en desacuerdo	En desacuerdo	De acuerdo	Muy de acuerdo
3.1. Mis padres están pendientes de mi rendimiento escolar.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3.2 Mis padres frecuentemente conversan con padres de mis compañeros, respecto a asuntos de mi Colegio.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3.3 Mis padres se aseguran que yo haga mi tarea y estudie para los exámenes.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

	Muy en desacuerdo	En desacuerdo	De acuerdo	Muy de acuerdo
3.4 Si tengo dificultades con una tarea, pido ayuda a mis padres.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3.5 Si mis calificaciones bajan, mis padres se aseguran que estudie más.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3.6 Mis padres se preocupan si no mantengo una buena disciplina en el Colegio	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3.7 La directiva de alumnos de mi curso fue elegida democráticamente	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3.8 Yo colaboro con las actividades de la directiva de alumnos de mi curso.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3.9 Nuestro profesor jefe apoya a que el curso plantee sus inquietudes u opiniones	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3.10 Estoy informado sobre lo que hace el Consejo de Alumnos Delegados del Colegio.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3.11 El Consejo de Alumnos Delegados del Colegio fue elegido de forma democrática.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3.12 El Consejo de Alumnos Delegados del Colegio recoge las opiniones de la mayoría de los estudiantes.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3.13 El Consejo de Alumnos Delegados es activo.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3.14 Las autoridades del Colegio le dan importancia a los planteamientos u opiniones del Consejo de Alumnos Delegados.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

4. Por favor señala tu nivel de participación en los siguientes organismos o actividades de tu Colegio (marca lo que corresponda frente a cada actividad)

	Nunca	Rara vez	Algunas veces	Siempre	No existe esta actividad
4.1 Consejo Escolar	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
4.2 Equipo deportivo del Colegio	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
4.3 Talleres extra-programáticos	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
4.4 Acciones solidarias	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
4.5 Grupo religioso	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
4.6 Grupo musical (banda, coro, etc.)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
4.7 Grupo cultural-artístico	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

5. ¿Cuán de acuerdo estás con las siguientes afirmaciones?

	Muy en desacuerdo	En desacuerdo	De acuerdo	Muy de acuerdo
5.1 Mi Colegio está muy comprometido con la promoción de valores morales.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
5.2 Siento que la mayoría de los profesores se preocupa por enseñarnos actitudes y valores importantes para nuestra vida futura.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
5.3 Con frecuencia hacemos trabajos grupales en el Colegio.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

	Muy en desacuerdo	En desacuerdo	De acuerdo	Muy de acuerdo
5.4 Las veces que he formado parte de un equipo, para trabajos del Colegio, ha resultado ser una buena experiencia.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
5.5 En este Colegio se respeta la diversidad.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
5.6 Los profesores en este Colegio permiten que los alumnos se "copien" trabajos o pruebas con nota.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
5.7 Mis profesores promueven la creatividad de los estudiantes.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
5.8 Los profesores estimulan el pensamiento crítico de los estudiantes.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
5.9 Los profesores estimulan el trabajo de investigación por parte de los estudiantes.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
5.10 Creo que mi Colegio funciona como una democracia.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
5.11 Creo que la educación que recibo en el Colegio me ayuda a desarrollar mi propia opinión frente a muchas situaciones.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
5.12 Por parte del Colegio participo en actividades de solidaridad y ayuda a otras personas.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
5.13 En el Colegio, consistentemente, hemos recibido orientación y educación sexual.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
5.14 En el Colegio, consistentemente, se nos ha enseñado a cuidar el medio ambiente.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
5.15 Siento que los profesores reconocen mis virtudes.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
5.16 Siento que, el Colegio se preocupa por estimular mi desarrollo como persona.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
5.17 En el Colegio se nos estimula a argumentar o fundamentar nuestras opiniones.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
5.18 En el Colegio se nos ha enseñado consistentemente que debemos considerar las consecuencias de nuestras acciones.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
5.19 En el Colegio se nos ha enseñado consistentemente a explorar distintas opciones antes de tomar una decisión importante en nuestras vidas.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

6. Señala tu grado de acuerdo con las siguientes afirmaciones sobre tu Colegio (marca lo que corresponda frente a cada afirmación)

	Muy en desacuerdo	En desacuerdo	De acuerdo	Muy de acuerdo	No lo conozco o no sé
6.1 Conozco las normas de convivencia del Colegio	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
6.2 Las normas de convivencia son respetadas por la mayoría de los estudiantes	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

	Muy en desacuerdo	En desacuerdo	De acuerdo	Muy de acuerdo	No lo conozco o no sé
6.3 Las normas de convivencia son respetadas por la mayoría de los profesores	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
6.4 Los alumnos hemos participado en la definición de las normas de convivencia del Colegio	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
6.5 En casos de faltas a las normas se aplican las sanciones establecidas en el reglamento de convivencia	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
6.6 Las sanciones se aplican por igual, sin discriminación a todos los estudiantes	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
6.7 En mi curso hemos elaborado normas de convivencia para nuestra aula.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
6.8 Cuando hago algo realmente bueno en el Colegio, mis profesores lo reconocen.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
6.9 Si estoy teniendo problemas personales, puedo conversar con mis profesores al respecto.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
6.10 Los profesores se ven poco motivados en su trabajo escolar	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
6.11 Los profesores de este Colegio motivan a los estudiantes para que logren su mejor desempeño.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
6.12 Los profesores de este Colegio me hacen sentir que soy capaz de realizar un buen trabajo.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
6.13 Si yo no entiendo algo, mis profesores tratarán de explicármelo de un modo diferente para que logre comprender.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
6.14 En general, las clases son entretenidas.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
6.15 Los profesores nos preguntan nuestra opinión durante las clases.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
6.16 Tengo temor de cometer errores en las clases.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
6.17 Los profesores tratan de relacionar el material de su clase con lo que nosotros estamos aprendiendo en otras clases	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
6.18 Los profesores animan a los estudiantes para participar activamente en las clases.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

7. ¿Con que frecuencia has observado las siguientes conductas por parte de...

7.1 ... los profesores de tu Colegio acostumbran?	Nunca	Pocas veces	A menudo	Siempre
7.1.1 Intimidar con amenazas a estudiantes.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

7.1.2 Golpear a estudiantes.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
7.1.3 Insultar a estudiantes.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
7.1.4 Ridiculizar a estudiantes	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
7.2. ... los estudiantes de tu Colegio acostumbran?				
7.2.1 Agredir a los profesores	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
7.2.3 Faltar el respeto a los profesores	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
7.2.3 Tener un comportamiento en la sala que impide hacer clases.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

8. Cuándo ocurre un conflicto en el Colegio ¿cuáles son las maneras más comunes de resolverlo?
(Marca solo 2 respuestas)

Dialogando y llegando a acuerdos	<input type="checkbox"/>
Con castigo y sanciones	<input type="checkbox"/>
Sermoneando a quienes se estima son los culpables	<input type="checkbox"/>
Tratando el conflicto entre todos: orientador, coordinador, profesor jefe, alumnos	<input type="checkbox"/>
No dándole importancia y dejándolo pasar	<input type="checkbox"/>
Buscando la ayuda de un mediador.	<input type="checkbox"/>

9. Desde que comenzaron las clases, ¿cuántas veces dentro del Colegio...

	0 veces	1 vez	2 o 3 veces	4 o más
9.1 has sido víctima del robo o daño intencional de tu propiedad, mp3, ropa o libros?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
9.2 has sido víctima de rumores o mentiras dañinas sobre ti?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
9.3 has tenido miedo de que alguien te golpee?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
9.4 has sido víctima de burlas debido a tu aspecto o por tu manera de hablar?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
9.5 te han empujado, hecho a un lado con un empujón, fuiste golpeado o pateado por alguien que no bromeaba?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
9.6 has estado en una pelea física?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
9.7 te han ofrecido, vendido o dado alguna droga ilegal?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
9.8 te han amenazado o lastimado con un arma (una pistola, una navaja o cuchillo, o una macana o palo, etc.)?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
9.9 te has sentido acosado sexualmente y has sentido miedo por esa razón?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
9.10 has dañado intencionalmente la propiedad de la escuela?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
9.11 has llevado contigo un arma?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

10. Señala tu grado de acuerdo con las siguientes afirmaciones (marca lo que corresponda frente a cada afirmación)

	No se aplica	Muy en desacuerdo	En desacuerdo	De acuerdo	Muy de acuerdo
10.1 La relación con mis hermanos/as es buena.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
10.2 La relación con mis papás es buena.		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
10.3 La relación entre mis padres es buena.		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
10.4 Frecuentemente hay fuertes discusiones familiares.		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
10.5 Yo quisiera irme de la casa lo antes posible.		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
10.6 Hay un buen clima familiar.		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

¡Muchas gracias por tu colaboración!

