



2010 - Volumen 3, Número 3



<http://www.rinace.net/riee/numeros/vol3-num3.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

Editora:

Verónica González

Comité Directivo:

Marcela Gajardo (PREAL - Chile)
Sergio Martinic V. (PUC - Chile)
Carlos Pardo (ICFES - Colombia)
Margarita Poggi (IPE/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)
Léonor 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)
Rubén Klein (Fund. Cesgranrio – Brasil)
Luis Lizasoain (UPV - España)
Jorge Manzi (PUC - Chile)
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 (UAA - 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 2010



Presentación

- Luces y Sombras de la Evaluación de Políticas y Programas Educativos 3
Marcela Román

Artículos

- Políticas de Avaliação das Escolas em Portugal 9
Maria do Carmo Climaco
- La Evaluación y las Reformas Educativas en América Latina 30
Sergio Martinic
- La Nueva Naturaleza del Aprendizaje y de la Evaluación en el Contexto del Desarrollo Competencial, Retos Europeos en la Educación del Siglo XXI 44
Joan Mateo y Dimitris Vlachopoulos
- O Efeito da Capacitação Docente no Desempenho Dos Alunos: Uma Contribuição para A Formulação de Novas Políticas Públicas de Melhoria da Qualidade da Educação Basica 62
Nilma Santos Fontanive y Rubén Klein
- Cuatro Hipótesis sobre la Institucionalización de la Evaluación Educativa en América Latina: El Uruguay en Mirada Comparada 90
Tabaré Fernández
- Estudio Diferencial del Rendimiento Académico en Lengua Española de Estudiantes de Educación Secundaria de Baja California (México) 115
Luis Lizasoain Hernández y Luis Joaristi Olariaga
- Problemas de Equidad en el Sistema Educativo Hondureño 135
Mario Alas Solís y Germán E. Moncada Godoy

Reseña

- Stake, Robert E. (2010). *Investigación Cualitativa: El estudio de cómo funcionan las cosas*. New York: The Guilford Press 152
Zulma Perassi



ISSN: 1989-0397

LUCES Y SOMBRAS DE LA EVALUACIÓN DE POLÍTICAS Y PROGRAMAS EDUCATIVOS

LIGHTS AND SHADES OF THE EVALUATION OF POLICIES AND PROGRAMS IN EDUCATION

Marcela Román

Revista Iberoamericana de Evaluación Educativa 2010 - Volumen 3, Número 3

<http://www.rinace.net/riee/numeros/vol3-num3/editorial.pdf>

Al despuntar la década de los noventa y con la llegada de gobiernos democráticos, la mayoría de los países de América Latina y el Caribe implementaron importantes procesos de Reforma Educativa. Desde sus inicios, estos procesos fueron acompañados con estudios y evaluaciones que analizaban desde diferentes matrices y miradas los efectos y cambios acontecidos en los sistemas educativos, escuelas y desempeños escolares producto de sus acciones, regulaciones, insumos y programas componentes. Hacia mediados de esa década los sistemas disponían ya de un interesante acumulado de evaluaciones sobre las políticas y programas que delinearon y configuraron el escenario educativo, sus propuestas y cambios. En un principio, en tanto exigencia directa de los organismos internacionales, tales como el Banco Mundial y el Banco Interamericano de Desarrollo (BID), quienes apoyaron financieramente a los distintos Estados para el diseño e implementación de tales innovaciones y cambios. La evaluación se convertía así (y aún lo es), en el principal mecanismo para rendir cuenta de lo invertido y logrado en términos de calidad y equidad de los sistemas, eje fundamental de las reformas educativas de la Región. Poco a poco la evaluación de programas y proyectos educativos se ha ido incorporando también a la cultura de investigadores, autoridades técnicas y políticas. Hoy no es posible pensar cambios o implementación de políticas, sin que ellas tengan evaluaciones o estudios pilotos previos, tanto como análisis de sus efectos en el mediano y largo plazo que permitan aventurar el cambio o la sostenibilidad buscada. Sin duda la evaluación de programas y, especialmente de los aprendizajes escolares, es el campo educativo de mayor crecimiento, expansión y sofisticación técnica de estas últimas dos décadas.

La evaluación de políticas, proyectos y programas en el campo educativo se constituye esencialmente en el juicio emitido sobre la estructura, el funcionamiento y los resultados de estas intervenciones educativas. Juicio que, aunque construido mediante un riguroso proceso investigativo y sostenido en ciertos principios y criterios preestablecidos, siempre será interpretativo y por ende discutible. A través de esta mirada, se describen y analizan así, etapas, acciones, estrategias, productos, efectos y distintos tipos de resultados implicados en la ejecución de dichos programas y proyectos. Ello con el propósito de determinar, entre otros aspectos, la pertinencia y logro de objetivos, la eficiencia, el impacto y/o la sustentabilidad de las acciones en el tiempo y para otras poblaciones similares.

Saber qué aprenden, qué son capaces de hacer y cuánto logran los niños, las niñas y los jóvenes estudiantes, se consolidan como aspectos esenciales y prioritarios desde donde la evaluación ha monitoreado la calidad educativa ofrecida en cada sistema y país de la región. Aunque menos visible y con menor debate técnico y escándalo público, se han evaluado también y, en más de una ocasión, las principales políticas y programas con distintos propósitos y por distintos agentes. Así por ejemplo, han sido sometido al escrutinio y rigor de la evaluación la casi totalidad de los programas focalizados emblemáticos: Escuela Nueva en Colombia, Programa de Mejoramiento de la Calidad de la Educación en las Escuelas Básicas de Sectores Pobres -P900 de Chile, EDUCO de El Salvador o Bolsas Escolares en Brasil, para mencionar sólo algunos. Las diversas evaluaciones que se han realizado a las propuestas pedagógicas de Fe y Alegría en los distintos países donde ella actúa, es otro ejemplo de evaluación de estrategias de focalización, pero desde la Sociedad Civil. Han sido evaluados también las políticas de supervisión, formación docente y fortalecimiento directivo, así como diversas iniciativas y estrategias más nuevas e integrales que han buscado el cambio y la mejora hacia escuelas y sistemas eficaces. Entre ellas destacan por su creciente protagonismo,

las asistencias técnicas externas a cargo de entidades y organismos privados o en alianzas público-privadas.

Estas evaluaciones, así como las comparaciones y balances entre países y la propia región, han sido desarrolladas por organismos técnicos internacionales, centros de investigación, institutos y universidades. Destacan entre los primeros, la OEA, PNUD, UNICEF, CEPAL, La Facultad Latinoamericana de Ciencias Sociales, FLACSO, el IPE UNESCO y el Programa de Educación para América Latina (PREAL). Entre los Institutos y centros de investigación y estudios a nivel nacional, es posible mencionar en primer lugar a los propios Sistemas Nacionales de Evaluación de la Calidad; a la Asociación Brasileña de Evaluación Educativa, Abave, (Brasil); al Instituto Central de Ciencias Pedagógicas (ICCP) de Cuba; al Grupo de Análisis para el Desarrollo (GRADE) de Perú; al Centro de Investigación y Desarrollo de la Educación, CIDE, en Chile, al Centro de Investigaciones Culturales y Educativas (CICE) de Venezuela; Instituto Colombiano para la Evaluación de la Educación, ICFE de Colombia; al Centro Nacional de Investigaciones Sociales y Educativas (CENAISE) de Ecuador, al Instituto Nacional para la Evaluación de la Educación, INEE (México, entre tantos otros.

Desde todos ellos, así como desde otras entidades similares, se han realizado importantes investigaciones, mediciones y evaluaciones de los distintos programas y componentes que han acompañado la implementación de las Reformas, en sus distintos momentos, contextos y países ya sea por iniciativa propia, o respondiendo a peticiones y demandas desde los propios Ministerios de Educación o entidades financieras internacionales. A pesar de los innegables aportes de esta línea en evaluación, son del todo escasas las publicaciones que dan cuenta de los resultados que han arrojado las diversas evaluaciones de los tantos programas y proyectos mediante los cuales los estados y gobiernos han buscado mejorar la calidad y equidad de sus sistemas educativos. Ello debido fundamentalmente, al carácter reservado y de embargo que caracteriza la mayor parte de estos estudios evaluativos.

¿Quién podría discutir que las evaluaciones realizadas han sido pilar y referente para la toma de decisión en política educativa? Sin embargo, es del todo discutible y opinable su aporte al mejoramiento de la calidad y justa distribución de ella, en tanto fin o propósito esencial de tal práctica analítica. Ciertamente entre las principales funciones de la evaluación, está en poder ofrecer información que permita tomar buenas decisiones y, ello se facilita cuando se aporta además al conocimiento que fundamenta las políticas y programas que se busca evaluar. Así, con más o menos éxito, estos juicios y miradas han permitido no sólo tomar decisiones en relación con la administración o desarrollo de estas políticas e intervenciones en el campo educativo, sino que han contribuido, de manera importante, al conocimiento que las fundamenta y define. Desde ellas se han podido ampliar y enriquecer las perspectivas conceptuales y estrategias de implementación de tales iniciativas, focalizar acciones y controlar de mejor manera las relaciones, dimensiones y factores que inciden en los resultados o en los cambios que se promueven y esperan lograr. La evaluación de un programa o proyecto, será siempre un proceso de construcción de conocimiento acerca del objeto o problemática educativa evaluada, al mismo tiempo que un espacio de reflexión y puesta a prueba de las hipótesis desde donde el programa o proyecto asume dicha problemática y ofrece soluciones.

Tal como lo hemos venido repitiendo en otros escritos, la evaluación es un arma poderosa en educación. Al mismo tiempo que desde sus potencialidades se le respeta y concibe como un aliado en el avance y mejoramiento educativo, se le mira con desconfianza y temor. Su rigor metodológico y creciente sofisticación técnica, hace que sea muy difícil contra argumentar con su evidencia. Su voz, que siempre retumba fuerte en el escenario educativo y social, casi siempre anuncia malas noticias por sobre orientaciones y caminos mejores por donde avanzar. Por eso se le teme. Desde ella se desnudan y reflejan especialmente las debilidades e ineficiencias, con las inevitables sanciones sociales y consecuencias políticas, profesionales e institucionales que ello implica y supone.

No obstante ello, es necesario insistir una y otra vez en evaluamos para aprender y mejorar y no solo para sancionar o encontrar culpables. La mirada evaluativa es rica en posibilidades y herramientas conceptuales y metodológicas para analizar y emitir juicios sobre ciertas iniciativas o políticas, tanto desde sus diseñadores e implementadores como desde sus beneficiarios o destinatarios. Su atenta mirada, visibiliza y deja expuestas las hipótesis y supuestos de las reformas y decisiones técnico políticas, mostrando flaquezas y fortalezas o lo adecuado o riesgoso de sus negociaciones y consensos, pero al mismo tiempo puede y debe entregar elementos, criterios y orientaciones que permitan corregir aquello que se ha mostrado débil e ineficiente, tanto como fortalecer y replicar las prácticas y estrategias que están en la senda adecuada.

Aprovechemos en tal sentido, la interesante diversidad de tipos y modelos de evaluación de programas y proyectos en el campo educativo: cuantitativas y cualitativas de acuerdo con el enfoque o referente metodológico; ex-ante, intermedia y ex-post, según el momento de su realización; internas o externas, en relación a la perspectiva del observador o, también evaluaciones de diseño, proceso, efectividad e impacto, en función del foco u objeto de dicha evaluación. Por último, y si de costos se trata, el campo distingue entre evaluaciones costo-beneficio y costo-eficiencia.

El potencial es enorme. La evaluación de proyectos y programas en educación, combina como pocos procesos investigativos, la voz y percepción de los actores, con la acción y práctica de usuarios y beneficiarios; la reflexión y análisis sobre trayectorias y procesos, con la identificación y cuantificación de productos y resultados; lo cultural o simbólico, con lo económico y estructural; la efectividad con la eficacia o la sustentabilidad; los diseños, con los impactos o los recursos y el tiempo con lo posible y logrado. Probablemente debido a ello, esta práctica analítica se ha constituido en objeto de estudio y conocimiento específico y propio. Así, se discuten, analizan y comparan las bondades de distintos tipos de evaluación en función del tipo de intervención educativa que se busca evaluar o la oportunidad de su desarrollo, tanto como la pertinencia y relevancia de sus enfoques y modelos para dimensionar y dar cuenta de objetivos, logros y resultados buscados.

Con este telón de fondo, la Revista Iberoamericana de Evaluación Educativa, dedica este número a la Evaluación de políticas, programas y proyectos en educación. En ella hemos intentado reunir los principales aportes y aprendizajes surgidos desde múltiples evaluaciones desarrolladas en distintas partes de Iberoamérica, desde distintas perspectivas conceptuales, paradigmas y estrategias metodológicas. El resultado ustedes los podrán apreciar al recorrer

la riqueza de sus textos, en donde investigadores y académicos de innegable prestigio comparten su mirada y experiencia en este campo analítico.

María do Carmo, doctora en Política Social y Educativa, comparte su rica y vasta experiencia en evaluación en un excelente texto que permite recorrer y dimensionar todo la dinámica, procesos y aprendizajes de las políticas de evaluación de escuelas en su país, Portugal, del cual fue Subinspectora General de Educación. El artículo enmarca las distintas iniciativas públicas y privadas de evaluación de escuelas en el contexto socio político en que ellas fueron emergiendo, permitiendo comprender sus características y aportes sobre el sistema y las propias escuelas. Dicho texto se detiene en el actual Sistema de Evaluación de Escuelas, próximo a completar cinco años y que se encuentra en un proceso de análisis y evaluación para ajustar el modelo, sus componentes y estrategias antes de iniciar un nuevo ciclo.

Comprender la evaluación, su objeto, estrategias y aportes desde los procesos de reformas educativas en América Latina, es la propuesta que nos ofrece Sergio Martinic, doctor en Sociología e investigador de la Universidad Católica de Chile. El artículo analiza las características de las reformas educativas en los últimos 25 años en la región, dejando en evidencia su incidencia en el objeto, método y uso de los procesos evaluativos que las han acompañado. El recorrido se detiene también en los cambios sociales que han afectado directamente a las reformas y a la organización de los sistemas educativos, entregando a la evaluación un nuevo lugar y otras funciones, no exentas de tensiones teóricas y metodológicas que ponen importantes desafíos a los países para producir y usar el conocimiento que emerge desde dichas evaluaciones para mejorar los procesos y resultados educativos.

Joan Mateo, Catedrático de Métodos de Investigación y Diagnóstico en Educación de la Facultad de Pedagogía de la Universidad de Barcelona y de amplia experiencia en medición y evaluación educativa, junto con Dimitris Vlachopoulos, investigador de la Universidad Autónoma de Barcelona, profundizan en el desafío que le pone a la formación y a la evaluación, la nueva naturaleza del aprendizaje sostenida desde el enfoque y desarrollo de competencias. En un sólido texto, los autores presentan la respuesta europea a esta problemática y desafío, dando cuenta del tránsito desde los modelos educativos basados en la transmisión del conocimiento a los centrados en el desarrollo competencial. Desde esta mirada, los autores sostienen que *“toda actividad de aprendizaje debe ser evaluada y toda actividad evaluativa no es sino una actividad de aprendizaje de lujo”*. Desde allí entonces, el lugar central que ocupa la evaluación en la reflexión y el análisis en el texto y sus conclusiones.

El análisis de los efectos de la capacitación docente en el desempeño de los alumnos y su aporte al diseño de nuevas Políticas Públicas que buscan mejorar la Calidad de la Educación Básica, se constituye en el núcleo de la atención y reflexión del artículo de Nilma Santos y Rubén Klein, coordinadora de Evaluación en Larga Escala y Consultor respectivamente, de la Fundación Cesgranrio en Brasil. El texto entrega evidencias y argumentos que permiten explicar el mejoramiento de los aprendizajes de estudiantes brasileños de primaria, a partir de la mejora de la calidad de los profesores de Lengua portuguesa y Matemática que participaron de un riguroso Programa de Capacitación entre los años 2003 al 2008. Interesante artículo que muestra cómo procesos formativos largos y sistemáticos, que atienden la calidad profesional de los docentes y se acompañan de evaluaciones periódicas

sobre el progreso de los estudiantes, afectan positivamente los aprendizajes y desempeños escolares.

Tabaré Fernández, profesor de la Universidad de la República del Uruguay y ex coordinador del Programa PISA en dicho país, ofrece cuatro desafiantes hipótesis para reflexionar sobre el proceso de institucionalización de la evaluación educativa en América Latina y desde allí analizar lo acontecido en Uruguay. Supuestos de profundizan en las fortalezas y debilidades de tales procesos, a partir de las reformas de segunda generación, la orientación liberal de los partidos de gobierno, la utilidad pedagógica de las pruebas de medición de aprendizajes y del control ejercido por las agencia encargadas de la evaluación.

La elaboración y validación de un modelo jerárquico lineal para dar cuenta del rendimiento académico en Lengua Española de estudiantes de Educación Secundaria mexicanos, es el foco y propuesta del artículo de Luis Lizasoain y Luis Joaristi, ambos profesores e investigadores de la Universidad del País Vasco/Euskal Herriko Unibertsitatea (España). En un riguroso y claro recorrido, los autores avanzan hacia un estudio diferencial de subgrupos extremos de centros y estudiantes. Aportan con evidencia para sostener que dentro de las variables que se pueden considerar como clásicas predictoras del rendimiento académico, hay algunas que inciden sobre el rendimiento de todos los estudiantes y centros mientras que otras sólo afectan a determinados subgrupos de la población. Sus hallazgos señalan así, la necesidad de contar con modelos específicos para los estudiantes más desfavorecidos debido a que las enormes diferencias socioeconómicas y culturales modifican, en este grupo, los predictores y las variables relevantes.

Por último, Mario Alas y Germán Moncada, docentes e investigadores de Universidad Pedagógica Nacional Francisco Morazán de Honduras, comparten una panorámica sobre las desigualdades y brechas de equidad del sistema educativo hondureño. Sus análisis profundizan en las diferencia de acceso, así como de los niveles de aprendizaje alcanzados por los estudiantes según características socioculturales, geográficas y de género. Desde allí, se concluye que la propia lógica de operación del sistema es responsable también de reproducir y acentuar las ya importantes brechas existentes.

Cierra la revista, el comentario de Zulma Perazzi, investigadora de la U. de Buenos Aires, Argentina, al reciente libro de Robert Stake (2010), dedicado a la Investigación Cualitativa. Stake, reconocido experto en evaluación educativa y Director del Centro de Investigación Instruccional y Evaluación Curricular en la Universidad de Illinois (USA), retoma la característica interpretativa de la investigación, para relevar la importancia que adquiere la habilidad de observar, definir y redefinir el sentido y significado de lo percibido.

Como siempre, les invitamos a leer cada uno de estos artículos y a reflexionar desde la información y evidencia que ellos ofrecen, respecto de cuánto han aportado las evaluaciones en mejorar la pertinencia y relevancia de lo enseñado en nuestras escuelas y liceos, o para dotar de mayor justicia el campo educativo. Discutir y debatir si la evaluación y los evaluadores, son también responsables de los pendientes, brechas y desigualdades que lamentablemente aún exhibe la mayoría de nuestros sistemas educativos.

POLÍTICAS DE AVALIAÇÃO DAS ESCOLAS EM PORTUGAL

EVALUATION POLICIES IN THE SCHOOLS OF PORTUGAL

Maria do Carmo Clímaco

Revista Iberoamericana de Evaluación Educativa 2010 - Volumen 3, Número 3

<http://www.rinace.net/riee/numeros/vol3-num3/art1.pdf>

Fecha de recepción: 13 de octubre de 2010
Fecha de dictaminación: 27 octubre de 2010
Fecha de aceptación: 30 octubre de 2010

1. AVALIAÇÃO DAS ESCOLAS EM PORTUGAL - CONTEXTOS E INICIATIVAS

A questão política e técnica da avaliação das escolas aparece de forma recorrente em Portugal em meados dos anos 80 e decorre essencialmente dos efeitos da contaminação das políticas derivadas da progressiva internacionalização do país (Barroso, 2006), quer por iniciativa dos serviços centrais da administração educativa e da sua interpretação dos objectivos dos programas dos governos, no caso das escolas públicas, quer por iniciativa de agentes privados associados a entidades académicas ou envolvidos em redes internacionais de estabelecimentos de ensino, no caso das escolas privadas.

Independentemente da sua orientação política, foram muito poucos os programas de governo que, de forma explícita, inseriram a questão da avaliação das escolas como objectivo político ou como estratégia. Frequentemente, os programas dos Governos nos seus objectivos programáticos mencionam a questão da *"reforma e modernização educativa"*, apresentada como *"um desafio nacional"* e referem questões relacionadas com a qualidade da educação escolar, referindo um conjunto de prioridades de acção governativa, especificamente:

- a *"melhoria da qualidade da educação, actuando prioritariamente sobre a reforma curricular"*;
- a *"reforma da administração educacional, com ênfase claro no reforço da autonomia da escola"*;
- papel da escola como *"núcleo activo de promoção e alianças estratégicas entre os vários intervenientes da comunidade educativa"*;
- a necessidade de *"mais intensa participação da comunidade na gestão da escola, com especial relevo para os pais, os empregadores, as autarquias e instituições locais"*;
- *"desenvolvimento de mecanismos de informação e comunicação com os clientes do sistema, por forma a fomentar a transparência e as responsabilidades acrescidas da escola perante a comunidade"*.¹

Não sendo a avaliação das escolas um objectivo expresso neste programa, foi no entanto neste período que foram construídas as bases necessárias para que surgissem diferentes iniciativas quer na esfera das escolas públicas, quer na esfera das escolas privadas.

No diploma² que estabelece o enquadramento político e normativo da autonomia das escolas refere-se também que, *"o Ministério da Educação adoptará as estruturas e mecanismos normais adequados para proceder à avaliação sistemática da qualidade pedagógica e dos resultados educativos das escolas sujeitas ao regime definido no presente diploma"*

Este pensamento foi contemporâneo das iniciativas internacionais de revisão e aperfeiçoamento dos sistemas de informação, especificamente as que foram conduzidas pela OCDE e que tinham como objectivo rever os sistemas de informação na área da educação, desenvolver novos indicadores para analisar e avaliar o desenvolvimento e a qualidade dos sistemas educativos³, especificamente, adequa-

¹ In Programa de Governo, sendo ministro da Educação Roberto Carneiro (1987/1991)

² Decreto – lei 43/89, de 3 de Fevereiro

³ O projecto INES (Indicadores Internacionais dos Sistemas Educativos) foi lançado na Conferência Internacional de Washington em 1987 e tinha como objectivo planear um sistema de indicadores que pudessem ser utilizados na monitorização e na análise comparada dos sistemas educativos. Em 1989 foi possível acordar numa estrutura de indicadores em 4 áreas consideradas fundamentais na avaliação dos sistemas educativos, dando origem à criação de 4 redes ou equipas de trabalho focalizadas em domínios específicos: as aprendizagens dos alunos (rede A), a educação e o mercado de trabalho (rede B), as escolas e os processos escolares (rede C) e expectativas e atitudes dos parceiros sociais face à educação (rede D).

los à avaliação do desempenho das escolas. Em vários estados membros as equipas nacionais colaboradoras da Rede C fizeram reflectir as orientações do projecto internacional na selecção, organização e produção de indicadores específicos para análise do desempenho das escolas, com o objectivo de melhorar o sistema de informação e testar o potencial informativo e heurístico daqueles instrumentos, como aconteceu em Portugal.

Podem-se encontrar diferentes enquadramentos conceptuais para justificar a produção dos indicadores de desempenho dos sistemas educativos em cada uma das redes deste projecto internacional, das teorias das organizações às teorias do controlo e da regulação. No que se refere aos critérios de selecção dos indicadores para descrever e avaliar as escolas, na rede C (desempenho das escolas) era muito explícita a influência da investigação sobre a eficácia das escolas e a sua relação com as políticas de equidade. A discussão do que se entendia por “uma escola de qualidade” ou por “uma boa escola” fez parte de vários fóruns nacionais e internacionais criados durante os anos 90, quer sob a égide da OCDE quer da própria União Europeia. Os seus objectivos eram os de aprofundar as dimensões que estão em jogo na definição da qualidade da escola e saber até que ponto as perspectivas divergiam conforme os actores participantes e poder confrontar o conhecimento produzido pela investigação com as orientações das políticas dominantes.

A educação comparada conheceu um novo rigor metodológico e uma relevância política consideráveis. A internacionalização das políticas contaminou as soluções encontradas em cada país para avaliar as escolas, identificadas como a “caixa negra” do sistema (Scheerens, 1997), ou o nível crítico onde se processam todas as diferenças qualitativas dos sistemas educativos.

Neste contexto, quer no quadro das políticas da modernização e da descentralização da administração educativa em Portugal, quer no âmbito da criação de mecanismos melhoria da qualidade da educação escolar, foram desenvolvidos projectos de monitorização e avaliação de largos sectores da vida das escolas nas modalidades de auto-avaliação e de avaliação externa, como vinha acontecendo em vários países de um e outro lado do Atlântico, nuns casos com carácter de projectos experimentais, frequentemente mais consentâneos com as características de investigação - acção, noutros casos com carácter de intervenção de regulação, explicitamente de avaliação externa.

1.1. A auto-avaliação para todas as escolas PEPT-2000

No âmbito do Programa Educação para Todos – PEPT-2000, criado para dar corpo ao desiderato do XI Governo (1987/91) de tornar obrigatório o cumprimento universal da escolaridade básica de 9 anos para todos os alunos até ao horizonte do ano 2000, foi instituída, em 1991, uma componente obrigatória de auto avaliação para todas as escolas candidatas e apoiadas por este Programa, que se obrigavam a criar um “Observatório da Qualidade da escola” (OQE). Como material de apoio essencial para as escolas organizarem os seus dispositivos de recolha e análise de informação, foi distribuído às escolas um Guião organizador desse dispositivo, onde se apresentava uma selecção de 15 indicadores quantitativos e

A cada uma destas redes cabia analisar o tipo de informação disponível nos sistemas de informação estatística nacionais, analisar e garantir a qualidade da informação recolhida, rever a operacionalização dos indicadores já estabelecidos e estudar metodologias de cálculo de outras medidas que a investigação sobre a equidade e a qualidade dos processos educativos vinha a indicar como relevantes. No domínio da rede C, a selecção dos indicadores para a comparabilidade internacional distribuíam-se em quatro categorias: os recursos humanos, os processos de decisão, o currículo, e os processos de organização e ensino a nível de escola. Deste trabalho global resultou publicação de vários relatórios parcelares e os relatórios anuais do programa INES “Education at a Glance” e o volume que o acompanha “Education Policy Analysis”, focalizado na análise e comparação de áreas específicas da política educativa.

qualitativos, dos quais as escolas eram convidadas a seleccionar 10, conforme a sua adequação ao seu projecto educativo. No seu conjunto esta selecção de indicadores cobria aspectos como os relativos aos resultados escolares, à caracterização da população escolar tendo em conta as características socioeconómicas e educacionais das famílias, os recursos humanos e os recursos físicos e tecnológicos, o nível e tipo de participação na vida da escola por parte de todos os implicados (professores, alunos, famílias e pessoal não docente), a coesão do corpo docente e o nível de satisfação de todos os actores. O primeiro objectivo destas tarefas era levar as escolas a adquirir competências na área do tratamento e análise da sua própria informação e tornar consciente a importância da tomada de decisão baseada em conhecimento e não em opiniões individuais ou na norma descontextualizada. Por isso, para cada indicador se apresentava a sua definição, o processo e os instrumentos para recolha e tratamento de dados, bem como a justificação da sua relevância. Como princípios justificativos e orientadores da auto-avaliação em cada escola, sublinhava-se o seu contributo como condição de autonomia e de participação activa.

A recolha sistemática da informação tornava possível a sua análise longitudinal, para conhecer a evolução da escola. O tratamento e a análise partilhada da informação a nível de escola era considerada em si mesmo um potencial contributo para a mobilização dos actores e, desse modo, *"para a qualidade da escola, significando melhorar a sua capacidade de responder às necessidades de aprendizagem", ... "mobilizar as comunidades educativas em torno das suas escolas" e ... "introduzir uma reforma cultural na gestão escolar, tornando-a transparente e rigorosa, capaz de se planear de uma forma estratégica e de avaliar os seus próprios resultados, orientada por princípios de democraticidade e de eficácia"* (Guião Organizativo do OQE, p.7-8).

O formato de auto avaliação decorreu da convicção de que as políticas da melhoria da qualidade da educação precisam tanto do apoio e da vontade dos políticos e da administração educativa, como do esforço reflectido em cada escola sobre a capacidade de fazer melhor. Esse esforço implica auto conhecimento.

O "Observatório da Qualidade da Escola" foi um duro processo de aprendizagem para muitas escolas, que por tradição não tratavam nenhum dos dados com que alimentavam as estatísticas nacionais, nem tinham sentido a sua falta. A sua surpresa foi descobrir como podiam pensar a sua escola "lendo" o significado dos números em cada indicador e comparar-se com outras escolas semelhantes. Durante os 10 anos de vigência do PEPT-2000, o "Observatório" foi essencialmente um projecto voluntarista dos seus promotores e das mais de 1000 escolas abrangidas, que receberam o apoio executivo mínimo por parte da administração educativa, para que não fossem desenvolvidas dependências burocráticas, contrárias à autonomia que se procurava.

O facto de ser possível às escolas contratar um "consultor" ou "amigo crítico" que as assessorasse nesse processo contribuiu para produzir conhecimento e reflexão diversificados sobre o desempenho das escolas e sobre as questões que lhes estão associadas. Na sua esteira, foram realizados e divulgados vários trabalhos académicos, tendo como objecto de estudo a avaliação do desempenho das escolas nos termos propostos pelo projecto. Por arrastamento, foram desenvolvidos outros "objectos de estudo" paralelos e outras formas de abordar a análise da qualidade das escolas, numa perspectiva metodológica predominantemente qualitativa, orientada para a compreensão dos objectos de estudo e menos focalizada na sua descrição "objectiva" e estatística.

Como refere Canário (1995), o ponto de entrada mais comum para a análise e a avaliação do desempenho educativo, fez – se tradicionalmente a partir da análise do Sistema Educativo, visto como um conglomerado de escolas, recorrendo à informação estatística centralmente recolhida, de forma sistemática. Poder olhar de perto as estatísticas de cada escola em conjunto com outra informação de diferente natureza, e poder procurar perceber como se criavam singularidades dentro da escola vista não como unidade periférica, mas como um sistema em si mesmo, era ter acesso a “uma porta de entrada” para o estudo da escola que normalmente não estava facilitada a muitos investigadores e estudiosos, agora tornados amigos críticos de “escolas PEPT”.

1.2. As iniciativas no âmbito da “Gestão da Qualidade” e da “Excelência”

Foram contemporâneas daquele projecto outras modalidades de autoavaliação, sobretudo as inspiradas no modelo da EFQM (Fundação Europeia para a Gestão da Qualidade), como o modelo CAF, cujos princípios orientadores são inspirados pelas teorias da Gestão da Qualidade Total. Foram as escolas particulares e cooperativas que mais consistentemente utilizaram a sua metodologia e organizaram as bases para o seu desenvolvimento, recorrendo à assessoria de um especialista e investigador universitário. Em várias zonas do país, ainda que de forma menos sistemática, muitas escolas públicas seguiram o exemplo, por opção, utilizando toda a instrumentação que fora produzida e entretanto divulgada.

O modelo EFQM assenta num corpo teórico de oito princípios fundamentais da “excelência” na gestão das organizações e que a definem como sustentável, a partir da sua articulação num “ciclo de qualidade”, de forma sistemática, continuamente revistos e melhorados. Os princípios que definem a sustentabilidade do modelo são, em primeiro lugar, a orientação para os resultados, e para os atingir, a focalização nos clientes ou usuários, a liderança e solidez dos objectivos, a gestão por processos e factos, o desenvolvimento e participação das pessoas, a aprendizagem, a inovação e melhoria contínuas, o desenvolvimento de parcerias e a responsabilidade social.

A sua matriz teórica com raízes na gestão de empresas e o discurso específico que a caracteriza foi fácil de absorver pelas escolas que dispõem de autonomia quer na gestão do seu projecto educativo, quer na gestão de processos e recursos. Mas a sua importação para as escolas públicas, mesmo quando adaptado⁴, não se tornou fácil de acomodar à sua cultura de conformidade ou à burocracia profissional, pondo em evidência que para além da competência pedagógica, científica, social e ética, as escolas precisam de se desprender do seu discurso cristalizado, avesso ao discurso técnico da organização e gestão de recursos, ou da análise de resultados conduzida com rigor. Do mesmo modo, precisam da cultura da invenção e do risco para conseguirem ir mais longe no seu desenvolvimento e melhoria, respondendo aos seus contextos específicos. Os contextos têm de ser vistos como desafios e oportunidades de uma outra acção e não como justificativas do que não se consegue alcançar.

1.3. O Estado regulador e os novos modelos de prestação de contas

A par destes “exercícios” de modernização da administração pública, outras fontes de influência têm feito sentir-se no modo como se encara a urgência de melhorar os resultados da educação e, por arrastamento, a prestação do serviço educativo. Decorrente da internacionalização do país, políticos e

⁴ Estamos-nos a referir ao modelo CAF (Common Assessment Framework), decorrente da adaptação do modelo EFQM para a avaliação dos serviços públicos no espaço europeu.

técnicos têm conhecido mais de perto as políticas de promoção da qualidade das escolas em uso noutros países Europeus, que se tornam fontes de inspiração para as políticas de modernização educativa.

Por toda a Europa assiste-se à reconceptualização do papel do Estado no desenvolvimento dos sistemas educativos, decorrente quer da expansão de políticas de tendência neo-liberal na gestão educativa com a valorização da autonomia das escolas e da liberdade de organizar o ensino e de escolher a escola de acordo com valores laicos ou confessionais (Afonso, 2003), quer do reforço das políticas de regulação e do controlo, mesmo em países onde a autonomia das escolas e um poder forte dos professores parecia uma tradição consolidada.

A standardização dos exames, o uso de provas aferidas, a avaliação das escolas por entidades externas credíveis e a autoavaliação das escolas foram-se tornando obrigatórios e usados como elementos reguladores das práticas educativas em muitos dos países dominantes, estendendo-se aos mais periféricos. A publicitação dos resultados dos alunos e das listas ordenadas das escolas a partir desses resultados tornaram-se formas de prestação de contas da "qualidade educativa" dos estabelecimentos escolares em vários países europeus, cujos efeitos se fizeram sentir quer nas políticas de financiamento da educação, quer na crescente competitividade entre escolas para garantir a sua procura pelos melhores alunos, quer na invocação do direito parental a escolher a melhor escola para os seus educandos. As regras do mercado invadem o mundo das escolas, mas deixam marcas importantes no clima das escolas e na insatisfação profissional dos docentes, que é preciso ter em conta.

Em paralelo, a investigação sobre a qualidade das escolas e sobre as condições do seu funcionamento que podem esbater as diferenças entre alunos, decorrentes da sua origem sócio económica e cultural, vai invadindo o campo da decisão política e condicionando os modelos de avaliação dos alunos, das escolas, dos próprios professores individuais, cada vez mais tornada uma peça fundamental na gestão das suas carreiras.

A avaliação tornou-se um campo científico e político. Simultaneamente, tornou-se uma nova área profissional, com efeitos na constituição de grupos profissionais de avaliadores organizados em associações privadas prestadoras de serviços de avaliação, e com efeitos na alteração do perfil dos que geriam as escolas e dos que na administração pública, tradicionalmente, acompanhavam as escolas e asseguravam o seu controlo e regulação – os inspectores.

A constituição de uma associação das inspeções europeias em 1995, a SICI (Standing International Conference of Inspectorates), de que Portugal foi um dos promotores, veio a tornar-se um importante fórum de debate sobre políticas de inspeção, modelos de avaliação e de troca de experiências, incluindo programas de formação de inspectores.

No final da década de 90, a institucionalização da avaliação das escolas tornou-se um objectivo do Governo português, tendo levado à reorganização das áreas de intervenção inspectiva, e tornando o recrutamento de uma nova geração de inspectores uma prioridade. Aos tradicionais papéis de acompanhamento, controlo e verificação da conformidade que caracterizavam a actividade inspectiva, vem juntar-se a responsabilidade de avaliar as escolas, exigindo uma nova maneira de ser inspector, a começar por deixar de trabalhar como agente individual para passar a ser membro de uma equipa.

A avaliação externa das escolas, tal como foi confiada à Inspeção Geral da Educação, foi concebida como um programa a desenvolver-se em ciclos de seis anos, de modo a avaliar todas as escolas, quer fossem unidades de gestão singulares, ou unidades de gestão em agrupamento de escolas⁵

O modelo orientava-se por “*uma filosofia de avaliação global do desempenho das escolas*”, - uma “*avaliação integrada*” em oposição às actividades tradicionais de inspeções fragmentadas por áreas de funcionamento das escolas - e justificava-se pela obrigação de prestar contas do trabalho desenvolvido com as crianças e com os alunos, dos resultados alcançados, e da boa utilização dos recursos. Eram destinatários desta prestação de contas a administração educativa e as comunidades educativas locais, na convicção de que a qualidade das escolas e das experiências de aprendizagem dos alunos é uma responsabilidade colectiva.

“Não haverá responsabilidade de uns pelos outros se não houver informação, pelo que a recolha de informação sistemática de dados através da observação local do desempenho escolar feita por agentes especializados tem uma função estratégica indispensável” (IGE, 2000, p. 9)

A “*observação local*” incluía a avaliação e discussão dos níveis de sucesso dos alunos, em absoluto e ponderados pelos factores do contexto familiar disponíveis⁶, bem como a sua comparação com o valor esperado do sucesso em cada escola; a análise documental sobre a gestão e organização da escola, incluindo os documentos orientadores do projecto e da vida da escola; a observação da realização do ensino e das aprendizagens em sala de aula, dos processos de avaliação das aprendizagens, bem como a organização e realização dos apoios educativos; e, ainda, a observação do “*clima*” da escola, tendo em conta aspectos como o relacionamento interpessoal, a participação e a liderança.

A estratégia deste programa e as diferentes operações que o compunham foram cuidadosamente preparadas e apresentadas a cada escola antes da intervenção no terreno. Do mesmo modo, o encerramento de cada intervenção ocorria num prazo curto, tendo como objectivo fazer a apresentação oral e a discussão de uma versão preliminar do Relatório de Avaliação a ser enviado na sua versão escrita, posteriormente.

Não foi um processo isento de sobressaltos para avaliadores e avaliados, todos envolvidos num novo modelo de relacionamento, onde a verificação da conformidade normativa não era o objectivo.

Se a avaliação iria introduzir uma nova cultura na gestão e direcção das escolas, como acontecera na larga maioria dos países Europeus, a entrada na sala de aula e a observação dos modos como os alunos trabalhavam foi em muitos casos quase um “*choque cultural*”, apesar de serem as escolas a escolher as aulas a observar.

⁵ Para assegurar melhor acesso à educação e aos seus recursos, incluindo uma maior racionalização na gestão dos recursos, e para garantir a coerência e sequencialidade dos percursos escolares dos alunos, os jardins-de-infância (educação pré-escolar) e as escolas básicas foram integrados numa única unidade de gestão, o agrupamento, num determinado território educativo, cuja sede de gestão é assumida pela escola do 3º ciclo, o ciclo final da escolaridade obrigatória. Conforme as características da rede escolar e a distribuição da população escolar, vários agrupamentos integram também o ciclo secundário, ficando a escola sede do agrupamento instalada na escola secundária. Cada agrupamento é, assim, uma “*escola*” feita de muitas escolas dispersas por muitos sítios, com os mesmos órgãos de direcção e gestão. Neste texto usamos o conceito de escola para significar a unidade singular ou o agrupamento.

⁶ - Para ponderar os factores de contexto familiar utilizou-se a variável “*nível de carência económica*” medido pela “*elegibilidade para apoios económicos escolares*”.

Nos três anos de duração do projecto foram avaliadas 2039 escolas e observadas 600 salas de actividades na educação pré-escolar e 10.062 aulas e situações de trabalho de aprendizagem nos ciclos da educação básica e secundária.

Foram produzidos 3 relatórios nacionais relativos às intervenções realizadas nos anos lectivos de 1999 / 2000, 2000 /2001 e 2001 /2002.

O balanço de três anos consecutivos de avaliação integrada nas escolas dos diferentes níveis de ensino, permitiu caracterizar o funcionamento e a qualidade do desempenho das escolas de todos os níveis de ensino e o modo como se diferenciavam entre si e identificar pontos fortes e pontos fracos do sistema de escolas. Permitiu sobretudo que as salas de aulas deixassem de ser um enorme conjunto de “jardins secretos”.

A análise dos seis pontos fortes⁷ identificados, nesse balanço, mostra que se referem a um campo disperso de qualidade organizacional e relacional que se poderiam organizar em duas categorias: uma refere-se a alguns aspectos da gestão do trabalho das aprendizagens, à dimensão afectiva subjacente a esse trabalho, abrangendo professores e alunos, bem como à atitude face à mudança; a outra categoria refere-se à qualidade dos serviços de psicologia e orientação, sempre referidos em todos os relatórios como funcionando bem, desde que existam.

No que se refere aos 10 pontos fracos, concentram-se na área da gestão do sucesso e do trabalho em sala de aula (8 itens) e na dinâmica da escola (2 itens) referindo-se à participação na vida da escola e à articulação com outras escolas, através de projectos e parcerias.

Os relatórios de avaliação de escola eram documentos longos, escritos numa linguagem nem sempre acessível aos leitores interessados em saber o que se revelava na avaliação das escolas dos seus educandos ou do seu bairro. Como instrumento de prestação de contas eram complexos. Mas as escolas receberam bem a versão escrita daquilo que já sabiam que era a sua avaliação.

Mas não foi pacífica a divulgação dos resultados da avaliação de cada escola e os dados disponíveis não tranquilizaram a opinião pública: alguma comunicação social deu-lhes visibilidade, ao divulgar alguns pontos fortes e fracos por nível de ensino, e foi ao terreno colher dados que situassem em contexto o que o Relatório Nacional avaliava. Em simultâneo, pressionava para a divulgação pública de todos os relatórios de escola, tal como vinha pressionando para a “libertação” dos dados dos resultados dos exames nacionais do ensino secundário e das classificações internas das escolas, de modo a poder fazer a ordenação das escolas e a estabelecer comparações por níveis de ensino e por regiões geográficas, como forma “mais segura” de dar um retrato da qualidade das escolas, como se vinha fazendo em alguns países europeus.

Libertados os dados, a partir de 2001 a comunicação social começou a publicar as listas ordenadas de todas as escolas portuguesas, incluindo as públicas e as privadas, acompanhadas de relatos colhidos no local, que dessem uma visão rápida dos contextos, dos modos como cada escola se via e reagia à avaliação publicada, vários testemunhos na primeira pessoa a dar voz aos dirigentes das escolas, aos pais, aos professores, aos alunos, a investigadores convidados. Para o grande público esta era a

⁷ - Todos os itens são avaliados numa escala qualitativa de 4 níveis de A a D, correspondentes a Muito Bom (A), Bom (B), Suficiente (C) e Insuficiente (D). São considerados “pontos fortes” os itens observados com medianas de nível B e pelo menos 25% de avaliações de nível A. São considerados “pontos fracos” os itens observados com medianas de nível C.

“avaliação das escolas” que punha em evidência que as escolas particulares ocupavam quase todos os 20 primeiros lugares e que as escolas públicas tinham piores resultados e que se concentravam nas posições finais da tabela.

O programa do Governo que entrou em funções em 2002, sendo ministro David Justino (2002-2004), trouxe uma nova orientação política ao sector da Educação, com um discurso mais radical a sustentar princípios fundamentais como “*defender uma escola assente no respeito por valores, como o trabalho, a disciplina, a exigência, o rigor e a competência em busca da excelência*” (...) “*prossequindo metas ambiciosas aferidas internacionalmente e combatendo as assimetrias sociais e regionais*”.

Estes princípios, que se filiavam também no programa de reforma e modernização da administração pública, arrastada de governo para governo, haveriam de se tornar o ponto de partida e a bandeira para sustentar a crítica ao sistema escolar herdado dos governos anteriores e para introduzir novas políticas de avaliação de todos os implicados no processo educativo: dos agentes às instituições, dos currículos aos manuais, dos alunos aos diferentes funcionários:

“O Governo considera essencial introduzir a todos os níveis uma cultura de avaliação das instituições, dos docentes, dos funcionários e dos alunos, que tenha consequências no seu desempenho e no desenvolvimento organizacional, profissional e humano”.

“a avaliação do desempenho das escolas, com publicitação dos resultados e criação de um sistema de distinção do mérito e de apoio às que demonstrem maiores carências.”

De entre outros enunciados programáticos, destaca-se ainda o propósito, sempre repetido em vários programas de governos com diferentes orientações políticas, de rever o sistema de informação do ME, tornando-o um sistema útil à gestão e à avaliação, coerente e de fácil acesso para a administração e para os cidadãos.

Nos 4 anos de governação, entre 2002 a 2004, tomaram-se várias medidas no campo da avaliação: publicou-se a Lei 31/ 2002 de 20 de Dezembro, sobre a avaliação do sistema Educativo e do ensino não superior onde se especifica a forma como se estrutura a avaliação, os conteúdos a abranger e as entidades coordenadoras; reestruturou-se o Ministério da Educação, extinguindo e fundindo serviços; acabou-se com o sistema de avaliação integrada das escolas que os governos anteriores tinham criado; entregou-se definitivamente à comunicação social a publicitação dos resultados dos exames nacionais do final do ensino secundário, 12º ano, e do final do ensino básico, 9º ano, e das provas aferidas do 4º e 6º anos, respectivamente os anos terminais do 1º e do 2º ciclos da educação básica:

“Os resultados da avaliação das escolas e do sistema educativo, constantes de relatórios de análise integrada, contextualizada e comparada, devem ser divulgados com o objectivo de disponibilizar aos cidadãos em geral e às comunidades educativas em particular uma visão extensiva, actualizada, criticamente reflectida e comparada internacionalmente do sistema educativo português”. (Lei 31/2002, art.º 16º)

Este artigo consagrou em lei a prática de publicação de listas ordenadas das escolas que já se iniciara em 2001. Alguns jornais de maior circulação e prestígio, continuaram a formar parcerias com diferentes entidades, nomeadamente do ensino superior, para o tratamento dos dados disponibilizados e publicação dos resultados de exames e provas aferidas, com diferentes níveis de desagregação, em listas ordenadas das escolas. Em alguns casos, as separatas publicadas constituem cadernos de jornalismo de investigação organizados por equipas de reportagem que percorrem o país, captando retratos e testemunhos sobre a

qualidade vivida em diferentes escolas e contextos geográficos e sociais, comparando diferentes fontes de informação, dando voz aos actores do sistema de escolas.

O debate sobre a organização das listas ordenadas das escolas com base nos resultados dos exames como instrumento de avaliação continua em aberto: uns regozijam-se com a vitória de dez anos de publicação de rankings: falam das escolas como unidades periféricas do sistema de administração educativa; outros insurgem-se com a imagem injusta e incompleta que apresentam das escolas: pensam nas escolas como sistemas sociais e organizacionais onde se vive, se trabalha e se aprende, quer se trate de crianças, de jovens ou adultos. São dois objectos de análise e de estudo diferentes, qualquer dos dois maltratados por este método de ordenação e comparação.

2. "MAIS E MELHOR EDUCAÇÃO": ENRAIZAR A CULTURA E A PRÁTICA DE AVALIAÇÃO

Em 2005-2009, com o novo Governo entretanto criado, sendo Ministra da Educação Maria de Lurdes Rodrigues, as questões da avaliação voltam a ganhar um novo lugar na política educativa. Sob a epígrafe "Mais e melhor Educação", o programa deste Governo define cinco propósitos essenciais para a política educativa. Sublinhada a negro no original, a última asserção constitui uma espécie de "trave mestra" do programa:

"Enraizar em todas as dimensões do sistema de educação e formação a cultura e a prática da avaliação e da prestação de contas".

Como alvos e critérios das novas políticas de avaliação, especifica-se e destaca-se no texto original:

"Avaliação do desempenho dos alunos e do currículo nacional, avaliação dos educadores e professores, avaliação, segundo critérios de resultados, eficiência e equidade, das escolas e dos serviços técnicos que as apoiam."

"... O Governo estimulará a celebração de contratos de autonomia entre as escolas e a administração educativa..."

"...Ao mesmo tempo, lançará um programa nacional de avaliação das escolas básicas e secundárias que considere as dimensões fundamentais do seu trabalho e não se reduza a uma ordenação sumária e acrítica baseada unicamente em notas de alguns exames, potenciando um modelo que tenha em conta os padrões adoptados no âmbito da União Europeia"

"... A avaliação terá consequências, quer para premiar as boas escolas, quer para torná-las referências para toda a rede, quer para apoiar, nos seus planos de melhoria, as escolas com mais dificuldades".

2.1. A operacionalização do programa nacional de avaliação de escolas

Para dar corpo a este ponto do programa de governo, foi constituído um grupo de trabalho para estudar e definir os referenciais da auto-avaliação e da avaliação externa dos estabelecimentos de educação pré-escolar e dos ensinos básico e secundário, *"identificando um conjunto de variáveis e parâmetros comuns às práticas de auto-avaliação desenvolvidas nos últimos anos e sugerindo outros parâmetros que as escolas pudessem escolher em função dos seus projectos e das suas condições específicas"*. Para a avaliação externa foi igualmente pedida a definição dos respectivos referenciais, devendo ser tido em conta que *"deste processo deveriam resultar classificações claras dos estabelecimentos de educação e ensino e recomendações que permitissem preparar a celebração de contratos de autonomia"*.

Competia ainda a este grupo de trabalho testar o modelo e a sua metodologia numa amostra de escolas convidadas a candidatarem-se a avaliação externa.

Na construção deste novo programa que viria a ser chamado de "Avaliação Externa das Escolas" tiveram-se em conta diferentes abordagens à avaliação institucional em uso em diferentes países parceiros na União Europeia, o modo como têm sido apresentadas na literatura científica, e os respectivos impactos que têm sido identificados e recapitulou-se o que a Lei 31/2002 sobre a avaliação das escolas dispunha quanto à sua estrutura e enfoques.

De igual modo, recapitularam-se as diversas iniciativas realizadas neste campo em Portugal, sobretudo as que tiveram maior implantação, públicas e privadas, os instrumentos de trabalho e os pontos fortes e fracos identificados quer nos próprios sistemas de avaliação, quer no desempenho das escolas.

2.2. Objectivos, metodologia e instrumentos

Tendo em conta a realidade contextual, o grupo consensualizou a importância de clarificar os objectivos deste programa de avaliação: se por um lado, de acordo com o programa do Governo e com o disposto na Lei, a avaliação tem como objectivo "enraizar a cultura e a prática da prestação de contas", e se, por outro lado, "toda a avaliação é formativa" (de acordo com Hadji) a avaliação teria como objectivo também produzir informação de retorno às próprias unidades avaliadas que lhes permitisse assegurar a consistência dos processos e dos resultados educativos, para todos.

2.3. O quadro de referência para a avaliação externa

Com esse objectivo em mente, o grupo de trabalho consensualizou um quadro de referência para a avaliação externa, a ser dado a conhecer antecipadamente às escolas. Este visa definir detalhadamente o que se quer avaliar e em que termos se propõe fazê-lo. O seu desenvolvimento fez-se em volta de 5 domínios articulados, que se definem através de uma questão, ou pergunta, apresentada como o fio condutor da avaliação em cada domínio, sintetizando o que se pretende saber em cada um, do seguinte modo:

Domínios de avaliação externa

1. *Resultados*: Como conhece a escola os resultados dos seus alunos, quais são e o que faz para garanti-los?
2. *Prestação do serviço educativo*: Para obter esses resultados, que serviço educativo presta e como o presta?
3. *Organização e gestão escolar*: Como se organiza e é gerida a escola para prestar esse serviço educativo?
4. *Liderança*: que lideranças tem a escola e que visão e estratégia está por trás da organização e gestão?
5. *Capacidade de auto-regulação e progresso da escola*: Como garante a escola o controlo e melhoria deste processo?

São estes domínios que, em última instância, serão avaliados de acordo com uma tabela de classificações de quatro níveis (Muito Bom, Bom, Suficiente e Insuficiente) cujos critérios de aplicação foram definidos previamente em termos descritivos.

Cada um destes domínios é, por sua vez, suportado por um conjunto de factores que os operacionalizam.⁸

Para habilitar os avaliadores a desenvolver um trabalho de campo convergente, seguindo o mesmo critério, e para possibilitar às escolas conhecerem esses critérios e prepararem-se para a avaliação externa, foi estabelecido para cada factor um conjunto de perguntas com o propósito de ilustrar diferentes aspectos relevantes a ter em conta na observação e nos diálogos a estabelecer, a gerir conforme a situação e o contexto. Estas perguntas se davam o “mote” para o desenrolar das entrevistas, constituíam não uma listagem do que perguntar, mas um modo de ilustrar como interpelar os actores e assim saber o que conheciam das suas realidades e como as sabiam “ler”, quer em termos de contexto, quer em termos de processos e de resultados.

2.4. O quadro de referência para a auto-avaliação

No que se refere aos referenciais da auto-avaliação, foi decidido não recomendar um modelo específico de auto-avaliação de escola, por diversas razões: primeiro porque todos os modelos são dinâmicos e nesse sentido vão evoluindo com vista a adequarem-se às exigências de melhoria das próprias escolas, conforme a sua realidade e, depois, porque qualquer opção deveria resultar de um acordo entre os vários membros de uma comunidade educativa concreta, que actua dentro de um contexto próprio.

Apesar da diversidade dos modelos, existem aspectos que são transversais a todos, como:

- A análise detalhada dos resultados das aprendizagens dos alunos,
- O enfoque no trabalho desenvolvido em sala de aula, com destaque para o modo como os alunos trabalham,
- O clima e o ambiente educativos
- A organização e gestão da escola e dos seus recursos.

É também comum uma maior ênfase nas questões qualitativas do processo das aprendizagens dos alunos e na atenção aos aspectos estratégicos da optimização dos recursos e da melhoria dos resultados. Teve-se igualmente como preocupação não confinar a auto-avaliação à produção de dados, segundo um modelo fixo, mas aceitar diferentes modos de o fazer, enfatizando a importância de cada escola saber ler a sua informação, saber interpretá-la para poder agir e saber prestar contas aos diferentes parceiros internos e externos, demonstrando o que foi feito, e os resultados obtidos.

Sabia-se que existia nas escolas uma grande diversidade de experiências e estádios de desenvolvimento dos processos de auto-avaliação, uns sistemáticos e estruturados, outros não estruturados, muitas vezes constituídos por vários episódios de apresentação de resultados dos alunos e sua discussão, ou de apresentação de dados relativos à avaliação dos diversos projectos institucionais previstos nos respectivos Planos de actividade, como recomendado nas normas de orientação da organização escolar.

Tendo em conta esta realidade, o Grupo de Trabalho optou por pedir às escolas que preparassem um texto para a sua apresentação de acordo com uma estrutura comum que lhes foi enviada, assente nos seguintes campos de análise:

⁸ Ver Anexos 1 e 2

1. *Contexto e caracterização geral da escola* – Qual o impacto, no funcionamento da escola, das características sociológicas dos diferentes núcleos onde se insere?
2. *O projecto educativo* – Que prioridades lhe estão subjacentes e que objectivos se estabelecem para as áreas mais relevantes?
3. *A organização e gestão da escola* – Como caracteriza o funcionamento das várias estruturas e órgãos da escola, e como se assegura a qualidade e equidade educativas?
4. *Ligação à comunidade* – Que articulação com as famílias e com as autarquias e instituições locais?
5. *Clima e ambiente educativos* – Que importância lhes é atribuída, bem como à educação para a cidadania e à disciplina?
6. *Resultados* - Como é feita a análise do sucesso dos alunos e como se avaliam os resultados académicos e sociais e o seu impacto?
7. *Outros elementos* relevantes para a caracterização.

Aquele texto, bem como a sua apresentação oral na abertura do processo de avaliação da escola, podem ser entendidos como um interface entre a auto-avaliação já realizada e a avaliação externa que se segue. Por esta razão os dois quadros de referência organizam-se seguindo a mesma metodologia: para cada campo de análise, identificam-se diferentes aspectos/factores que constituem a estrutura da apresentação comum a todas as escolas. Para cada factor enunciam-se conjuntos de questões que tornam mais preciso o que se pretende saber e o modo de resposta. Para além da resposta directa, deve ainda ser demonstrada a reflexão que foi feita nas escolas e as conclusões a que chegaram.

2.5. O trabalho de campo

Antes de se iniciar o trabalho de campo no calendário previsto, o qual foi dado a conhecer a cada escola com bastante tempo de antecedência, cada equipa de avaliação analisou a documentação solicitada às escolas a avaliar. Para além do relatório de auto-avaliação e do texto de apresentação de escola nos termos propostos, as escolas enviavam os documentos fundamentais de orientação, planeamento e gestão da escola. A leitura desta documentação permite obter uma primeira imagem da instituição: conhecer o contexto social e cultural das famílias dos alunos das escolas do agrupamento, a evolução dos resultados escolares nos últimos três anos, as diferentes ofertas educativas organizadas pela escola previstas no seu plano anual de actividades e saber como tudo se articula com o projecto educativo do Agrupamento. Esta documentação permite formular um conjunto de questões sobre a sua organização e actividade, em articulação com as questões previstas nos documentos de apoio ao trabalho de campo, especificamente o “Quadro de Referência” da avaliação externa.

Durante 3 dias em cada agrupamento uma equipa de três avaliadores, dois inspectores e um avaliador externo (sem ligação à Inspeção Geral da Educação) recrutado predominantemente no ensino superior, desenvolvem o trabalho de campo recolhendo dados a partir das visitas às escolas, da observação directa e dos testemunhos de um largo conjunto de entrevistados em painel, abrangendo um leque variado de elementos da escola e das comunidades locais representando, entre outros, os órgãos de direcção e coordenação da escola, professores com cargos de gestão intermédia, professores sem cargos, pais, alunos, não docentes, representantes dos serviços locais relacionados com a escola – saúde, autoridades municipais, serviços especializados em apoios educativos.

2.6. O Relatório de avaliação

A avaliação externa de cada unidade de gestão conclui-se com o envio de um relatório elaborado pela equipa de avaliadores. Todos os Relatórios obedecem a uma matriz pré -definida, de onde constam:

1. *Uma introdução*, comum a todos os relatórios, apresentando a actividade.
2. *A caracterização da unidade de gestão*, contextualizando-a.
3. *As conclusões da avaliação* em cada domínio e as respectivas classificações.
4. *A apreciação dos factores* que descrevem ou operacionalizam cada domínio, especificando as evidências que suportam os diferentes juízos de valor e a classificação do domínio.
5. *As considerações finais*, sintetizando os “*pontos fortes*”, “*os pontos fracos*”, as “*oportunidades de desenvolvimento que constituem desafios para a melhoria da escola*”, e “*constrangimentos a ser resolvidos*”.⁹

Em documento interno, dirigido aos avaliadores, contendo as orientações para a sua elaboração, explicita-se a preocupação de que o Relatório “*seja simples, útil e de fácil exequibilidade, considerando os diferentes destinatários*”.

Cada relatório é enviado à respectiva escola, que, por lei, tem um prazo de 10 dias para apresentar o seu contraditório.

Todas as alterações introduzidas no texto por força do contraditório são assinaladas, apresentando a versão original e a versão corrigida.

Entre 2006/2007 e 2009/2010, foram elaborados, enviados às escolas e publicados no sítio da IGE, 984 relatórios e 314 contraditórios. Estes valores globais têm uma distribuição variável em cada um dos anos lectivos e em cada uma das cinco regiões em que se divide o território nacional. Não deixa de ser significativo que cerca de 32% das escolas tenham sentido necessidade de se pronunciarem em sede de contraditório, quanto à apreciação e ou classificação que foi feita.

2.7. A avaliação anual do processo

No final de cada ano é aplicado a todas as escolas avaliadas e respectivos avaliadores um inquérito por questionário, estruturado num conjunto de questões de resposta fechada e algumas de resposta aberta sobre o processo em si, incluindo os instrumentos utilizados, o modo como as escolas se prepararam para a avaliação, o trabalho dos avaliadores, o relatório de avaliação externa, aspectos e propostas a melhorar neste processo.

De um modo geral as respostas das escolas são muito positivas quer para as equipas de avaliação, quer para a Inspeção Geral de Educação, que, dizem, foram capazes de criar um ambiente de cordialidade e abertura que os surpreendeu.

A generalidade dos inquiridos concorda com os instrumentos utilizados e com o modo da sua aplicação, considerando-os facilitadores dos processos de auto-avaliação e de avaliação externa, mas põem em destaque um conjunto de aspectos omissos que seria importante apreciar no âmbito da avaliação das

⁹ Ver Anexo 3

escolas, mas que nem o tempo dedicado a esta intervenção, 3 dias, nem a concepção do modelo o permite fazer.

Na apreciação que fazem dos Relatórios individuais por escola verifica-se um grande número de referências negativas à “justiça das apreciações”, em consonância com o número de contraditórios recebidos. Referem-se às inúmeras coisas que não foram vistas e, num tom de discurso de protesto e agressivo, alguns afirmam “não nos revemos nesta apreciação”

Um largo número de escolas envia contraditórios apenas a identificar pequenas gralhas, ou propostas de uma redacção ligeiramente mais precisa para ser mais verdadeira. Outros contestam as classificações e as descrições de desempenho e referem que “não existe relação entre o descrito e as menções qualitativas atribuídas”, sublinham a não consistência de critérios que leva a que “textos semelhantes nos descritivos correspondam a menções qualitativas diferentes”, ou “a excessiva ênfase nos aspectos negativos”. Consideram que este “é um modelo subjectivo” e que é preciso evitar “ambiguidades na avaliação”. Outros argumentam que “não se podem utilizar os mesmos parâmetros em todas as escolas sem ter em conta a realidade de cada uma” que não vêm ter sido suficientemente ponderada e sublinham “o contexto” como condicionante a ter em conta na avaliação do desempenho.

Quanto à avaliação que os avaliadores fazem do processo de avaliação externa, mantém-se constante o seu modo de resposta positiva ao longo dos quatro anos de auto-avaliação do projecto. Referem-se, no entanto, a um conjunto de observações sobre aspectos a melhorar, nomeadamente o que se refere à informação de caracterização das escolas recolhida e tratada pelos serviços da administração, o tempo demasiado curto para a realização do trabalho de campo da avaliação e a operacionalização dos níveis da escala de classificação.

3. PREPARAR O NOVO CICLO DE AVALIAÇÃO: AJUSTAMENTOS E MELHORIAS

Em 2011 concluir-se-á o ciclo de avaliação de todas as escolas portuguesas. Parece ser consensual que se deve procurar estabilizar o modelo actual, devendo promover-se um debate alargado fundamentado em investigação sobre este processo, sobre os seus pontos fortes e fracos antes da sua revisão e aprofundamento.

A leitura dos relatórios e respectivos contraditórios, bem como os comentários de avaliadores e avaliados, e os estudos em curso, sugerem vários aspectos concretos que merecem ser reajustados, nomeadamente:

i) O quadro de referência para a avaliação externa é extenso e complexo. Será preciso voltar a analisar a sua matriz e identificar o que está a mais ou precisa de ser reescrito. Por exemplo, deve ser questionado se os factores seleccionados se mantêm com a mesma pertinência que nos pareceu quando foi feita a primeira redacção, ou se não haverá redundâncias e, portanto, sobreposição de áreas a avaliar.

Também não é seguro que haja exemplos certos de perguntas que interpelem quem trabalha nos jardins-de-infância ou nas escolas do 1º ciclo, tal como faltarão as questões que se dirijam às especificidades dos outros níveis de ensino.

ii) Ao privilegiar-se uma perspectiva generalista de agrupamento, não é seguro que através dos relatórios se fique a saber com clareza:

- Que níveis de ensino, ou que grupos de alunos, são considerados “críticos” nos seus processos de aprendizagem e percursos escolares;
- Como se garante a sustentabilidade para a melhoria: quem acompanha, apoia, capacita as escolas e os seus actores na sua acção e auto-regulação, de forma sistemática.
- O que deixou de ser feito por falta de autonomia.

iii) É importante referir que predominam os relatórios muito bem escritos, precisos na informação, concisos no texto e consistentes nos comentários e na respectiva avaliação. Tal decorre não só do modo como se descrevem e apreciam os factores, mas como se organiza a síntese por domínio e se atribui a respectiva classificação.

Mas é igualmente importante garantir que todos os destinatários dos relatórios percebem o que está escrito. O discurso tem de ser mais simples, menos técnico, menos orientado para os docentes e técnicos de educação, dentro de uma linguagem comum que todos os cidadãos entendam. Um dos objectivos da avaliação diz respeito ao envolvimento e responsabilização de todos os parceiros da comunidade local na vida da escola e no respectivo processo de avaliação. O relatório não pode ser um instrumento de exclusão.

iv) Apesar de alguns ajustamentos entretanto introduzidos nos instrumentos e nos procedimentos de avaliação externa, bem como do investimento feito na formação dos avaliadores, a leitura sistemática dos Relatórios de Avaliação sugere que outro tipo de ajustamentos deveriam ser realizados antes de se entrar num novo ciclo de intervenções em 2011/2012. Algumas debilidades dos relatórios, quer em termos de forma, quer de conteúdo, podem reflectir-se na credibilidade da própria avaliação. Seleccionámos, por isso, quatro aspectos que se consideram fundamentais para melhorar a função de prestação de contas e a função de feedback para potenciar a melhoria da escola, nomeadamente:

- Um tratamento mais consistente dos resultados educativos dos alunos, de modo a terem a relevância que merecem. No caso dos resultados académicos, na ausência de metodologias mais elaboradas como o cálculo do valor acrescentado, a ser desenvolvidas e disponibilizadas pelos serviços centrais de estatísticas e informação, é importante conhecer pelo menos a distribuição dos níveis de classificações alcançados nas diferentes disciplinas do currículo, ou o percurso de cohortes de alunos, e saber como se correlacionam com o estatuto familiar dos alunos. No caso dos resultados sociais, afectivos ou de atitudes e expectativas face à escola, as evidências apresentadas nem sempre são sustentadas por recolhas e análises rigorosas, mas em relatos impressionistas. Não quer isto dizer que não sejam verdadeiros, mas não são rigorosos e podiam sê-lo.
- O insuficiente tempo para o trabalho de campo. Este facto pode ter como efeito o recurso a um discurso defensivo por parte dos avaliadores, que dispõem essencialmente de um conjunto de impressões que recolheram nas suas observações e nos testemunhos dos entrevistados, mas que não tiveram tempo

de triangular para testar a sua consistência e formular um juízo de valor sustentado em evidências claras, seguras.

- A revisão dos modos de definição e tratamento dos níveis de classificação. O seu carácter descritivo e qualitativo não está isento de interpretações e aplicações subjectivas, tornando possível a atribuição de classificações diferentes a descritivos iguais de domínios observados. O mesmo acontece com a identificação dos pontos fortes e fracos apresentados nas considerações finais, que por vezes expressam opiniões e não juízos fundados em critérios escrutináveis.
- O conhecimento que todos na escola têm da sua auto-avaliação interna, como a trabalharam e o que fizeram com ela. Para isso é preciso reportar a divulgação que foi feita, como foi apresentada e discutida, as consequências que houve no planeamento geral dos departamentos e dos professores individualmente, na organização e gestão dos recursos e na consecução dos objectivos de curto e médio prazo enunciados. Isto é, como se articularam e coordenaram entre si nos processos para conseguir resultados melhores: como partilharam práticas e reflectiram acerca dos resultados quer de modo informal, quer formalmente, onde encontraram os apoios e as lideranças necessárias. Não são claras as dinâmicas suscitadas pela auto-avaliação.

Para além de ajustamentos deste tipo, é preciso, também:

- Compilar ou produzir material de apoio a usar em auto-formação de avaliadores e avaliados, que ajude a explicitar os aspectos menos claros e as mensagens “demasiado implícitas” dos dois quadros de referência utilizados.
- Promover, com recurso a diferentes entidades formadoras, oportunidades de formação para todos os avaliadores, recorrendo a actividades / exercícios práticos de avaliação, de análise e de redacção de relatórios.

Em conclusão, melhorar o que temos é um desafio maior e mais estimulante do que anular o que foi feito e começar de novo. Iniciar um novo ciclo é uma oportunidade de aprender e de melhorar. Por isso, importa querer saber o que a comunidade científica e as comunidades locais, avaliadores e avaliados pensam como se pode tornar os instrumentos de trabalho mais adequados. Mas é também uma oportunidade para rever os pontos fortes dos vários programas de avaliação das escolas em Portugal e as palavras-chave que foram as suas “bandeiras” ou princípios orientadores:

- *Garantir a qualidade* – ter informação e saber produzir conhecimento sobre a qualidade da escola.
- *Orientar-se para a excelência*, manifestada nos seus oito conceitos fundamentais e garantir um ciclo contínuo de qualidade.
- *Melhorar o desempenho de uma forma integrada*, a partir dos processos de trabalho em sala de aula.
- *Interpelar a consistência* dos processos seguidos e dos resultados alcançados para todos.

Eles mostram-nos que a avaliação perfeita não existe, mas vai-se construindo progressivamente. Mostram-nos, também, que nos últimos 20 anos a avaliação tornou-se um processo cada vez mais crítico, mais exigente, mais complexo, mais profissional, mais técnico, mas mais centrado nas pessoas e no seu desenvolvimento.

REFERENCIAS BIBLIOGRÁFICAS

- Afonso, N. (2003). A regulação da educação na Europa: do Estado educador ao controlo social da escola pública. *João Barroso (org) A escola Pública. Regulação, Desregulação, Privatização*. Edições ASA.
- Avaliação Externa das Escolas. Referentes e Documentos de Trabalho (2009). www.ige.min-edu.pt.
- Avaliação Externa das Escolas. Relatório 2008-2009, (2009). Lisboa: IGE–ME.
- Avaliação Integrada das Escolas. Relatório Nacional. ano lectivo 1999/00; 2000/2001. Lisboa: IGE–ME. 2001/2002 (edição policopiada).
- Avaliação Externa das Escolas. Relatório 2008-2009. www.ige.min-edu.pt.
- Avaliação Externa das Escola: Relatórios por Região e por unidade de gestão (2006/07; a 2009/2010). www.ige.min-edu.pt.
- Barroso, J. (2006). *A Regulação das políticas públicas de Educação*. Educa.
- Coelho, I., Sarrico, C., e Rosa, M.J. (2008). Avaliação das escolas em Portugal: que futuro. *Revista Portuguesa e Brasileira de Gestão*, Abril/Junho, 2008. www.scielo.oces.mctes.pt/pdf/rpbg/v7n2.
- CNE, (2005). Estudo sobre *Avaliação das Escolas: Fundamental Modelos e Operacionalizar Processos*. www.cne.pt.
- MacGilchrist, B. (2003). *Has school improvement passed its sell-by-date?*. Institute of Education: University of London.
- Mortimore, *et al.*, (1988). *School Matters – the junior years*. Open Books.
- Observatório da Qualidade da Escola, (1995). 2ª edição. ME: PEPT-2000.
- Oliveira, P.G., *et al.*, (2006). Relatório final da actividade do Grupo de Trabalho para a Avaliação das Escolas.
- Oliveira, J.M., Presa, J.L., Marques da Silva, J. e Pereira, M.L. (2010). *Parecer sobre Avaliação Externa das Escolas*. www.cne.pt.
- Programa de Governos Constitucionais. www.google.com.
- Sammons, P., Myers, K., and Mortimore, P. (1997). *Forging links*. Londres: P.C.P.
- Sarrico, C.C, e Rosa, M.J. (2008). *Escolas Portuguesas avaliadas em 2007 – Uma análise de valor acrescentado*. Universidade de Aveiro.
- Scheerens, J., Botanni, N., and Tuijnman, A. (org), (1995). *Measuring the quality of schools*. Paris: OCDE.
- Silva Marques, A.A. e Marques da Silva, (2008). *Parecer sobre Avaliação Externa das Escolas*, www.cne.pt.

Anexo 1

O Modelo de Avaliação Externa

Domínios de Avaliação

- Resultados
- A prestação do serviço educativo
- A organização e gestão escolar
- Liderança
- Capacidade de auto-regulação e progresso da escola



Anexo 2

Domínios e Factores de Avaliação

▶ Resultados

- Sucesso académico
- Participação e desenvolvimento cívico
- Comportamento e disciplina
- Valorização e impacto das aprendizagens

▶ A prestação do serviço educativo

- Articulação e sequencialidade
- Acompanhamento da prática lectiva em sala de aula
- Diferenciação e apoios
- Abrangência do currículo e valorização dos saberes e da aprendizagem

▶ A organização e gestão escolar

- Conceção, planeamento e desenvolvimento da actividade
- Gestão dos recursos humanos
- Gestão dos recursos materiais e financeiros
- Participação dos pais e outros elementos da comunidade educativa
- Equidade e justiça

▶ Liderança

- Visão e estratégia
- Motivação e empenho
- Abertura à inovação
- Parcerias, protocolos e projectos

▶ Capacidade de auto-regulação e melhoria da escola

- Auto-avaliação
- Sustentabilidade do progresso



Anexo 3

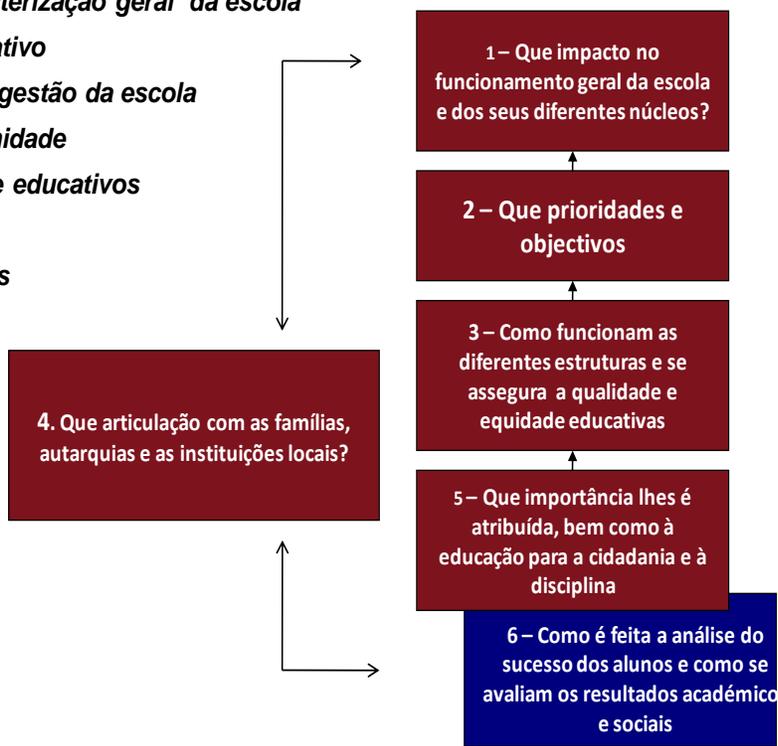
Campos de análise

O Modelo de Auto-Avaliação

➤ Contexto e caracterização geral da escola

- O projecto educativo
- A organização e gestão da escola
- Ligação à comunidade
- Clima e ambiente educativos
- Resultados

➤ Outros elementos



LA EVALUACIÓN Y LAS REFORMAS EDUCATIVAS EN AMÉRICA LATINA

THE EVALUATION AND THE EDUCATION REFORMS IN LATIN AMERICA

Sergio Martinic

Revista Iberoamericana de Evaluación Educativa 2010 - Volumen 3, Número 3

<http://www.rinace.net/riee/numeros/vol3-num3/art02.pdf>

Fecha de recepción: 15 de octubre de 2010
Fecha de dictaminación: 30 de octubre 2010
Fecha de aceptación: 30 octubre de 2010

En los últimos 20 años la evaluación educativa ha tenido un importante desarrollo en América Latina. Este desarrollo se debe a los procesos de Reforma Educativa que se llevan a cabo en la Región. Los países han entendido que el conocimiento e información producida a través de las evaluaciones es fundamental para dar cuenta de los logros; comprender los problemas y realizar los cambios necesarios para mejorar los procesos y resultados educativos.

La preocupación por la calidad y equidad de la educación de las Reformas ha ampliado los aspectos de la realidad que son objeto de las evaluaciones y ha diversificado las estrategias metodológicas. Tradicionalmente las evaluaciones se han concentrado en los aprendizajes curriculares, especialmente en matemática y lenguaje o lengua. Hoy día se evalúan además, y con distinto grado de desarrollo metodológico, la gestión directiva; los procesos educativos; los desempeños de profesores; el uso de materiales, tecnología y textos escolares, entre otros.

Las reformas educativas, a su vez, han transformado el centro de la decisión y se evoluciona en los países hacia sistemas descentralizados y participativos. En la región se vive el tránsito de una sociedad receptora de beneficios a otra más activa, con fuerte expresión de demandas y con más poder para ejercer control y participar en las decisiones de las unidades educativas. Todos estos cambios en la sociedad y la organización de los sistemas educativos inciden en el desarrollo de las evaluaciones otorgándoles, a la vez, un nuevo lugar y otras funciones.

En este artículo se analizará las características de las reformas educativas en los últimos 25 años y cómo estas inciden en el objeto, método y uso de los procesos evaluativos. Se analizarán las principales tensiones teóricas y metodológicas que plantea el desarrollo de la evaluación y se concluye con la identificación de algunos de los desafíos a resolver en el futuro.

1. CICLOS DE REFORMAS Y EVALUACIÓN

En los últimos 25 años en América Latina se han emprendido importantes Reformas Educativas. Siguiendo a Oscar Oslak (1996, 1999) es posible reconocer en dichas Reformas tres ciclos o generaciones de cambios. La primera de ellas -y que predominó en los años 80's- son reformas institucionales orientadas a reorganizar la gestión, financiamiento y acceso al sistema. Las segundas -propias de los 90's- están centrada en los procesos y resultados de los sistemas educativos y, por último, las reformas de tercera generación -propias de los 2000's- transforman las relaciones del estado con las escuelas y los modos de pensar, organizar y de poner en práctica los procesos educativos. En cada uno de estos ciclos es posible identificar paradigmas que predominan en la investigación y evaluación educativa y problemas particulares que se privilegian en función de las demandas y requerimientos del sistema.

Las reformas institucionales o de primera generación de los años 80 se dirigieron a transformar las estructuras de los servicios educativos. En la mayor parte de los países de América Latina se vivieron ajustes estructurales que, entre otros, implicó la descentralización de los servicios hacia los estados, las provincias, los municipios y las instituciones escolares. Se trató de reformas "hacia afuera" a través de las cuales los gobiernos centrales transfieren al sector privado y a los estados, provincias y comunas la gestión y administración de gran parte de los servicios educativos. En casi todos los países estos procesos estuvieron acompañados de una reducción del aparato público y de una racionalización de los recursos públicos del sector. Los énfasis de la política se concentraron en la accesibilidad y cobertura de la

educación básica; en el mejoramiento de la eficiencia y en la asignación de recursos según criterios de focalización.

En la década de los 90 se inician las reformas de "segunda generación". Estas tienen un movimiento contrario y se dirigen "hacia adentro". Es decir, hacia los modos de gestión y evaluación del sistema; los procesos pedagógicos y los contenidos culturales que se transmiten en la escuela. Se promueven políticas que otorgan mayor autonomía y poder a directores; se realizan cambios en el currículum; en la formación inicial y capacitación de profesores. Es un periodo de importantes inversiones en infraestructura, textos y otros insumos en las escuelas más pobres de los países de la región.

En la actualidad, se aprecian las señales un tercer ciclo de reformas y que redefinen las relaciones entre escuela y sociedad. Los cambios de tercera generación tienen unidades educativas más autónomas pero, a la vez, con mayor interconectividad y colaboración en red. Ponen el acento en la calidad de los resultados más que en las estrategias o caminos para lograrlos. Los establecimientos rinden cuenta y los actores tienen mecanismos y autoridad para exigirlos.

Es posible identificar algunos ejes para realizar la comparación de estos ciclos de reformas. En la tabla siguiente hemos relevado aquellos que nos parecen centrales para dar cuenta de los rasgos propios y comunes de los procesos de reforma en cada uno de los ciclos definidos.

En primer lugar, el eje o foco que interesa controlar desde la autoridad política ya que a este se atribuye poder central para el cambio. Observamos una evolución interesante desde el control de los recursos, preocupación central en la década de los 80, hacia el control de los resultados y que pone énfasis en la rendición de cuentas (Schleicher, 2006).

Se definen luego, actores, políticas ejes, y procesos claves y que aluden al núcleo organizador de las principales políticas de cada ciclo de reformas. En los 90, y por ejemplo, para el caso de Chile, hemos definido como política eje el cambio curricular. En efecto esta fue el corazón de la reforma y núcleo articulador de numerosos políticas y programas. Se promovió en el periodo un fuerte cambio en los contenidos de la educación esperando con ellos generar un nuevo lenguaje y condiciones para las prácticas pedagógicas.

Se compara a continuación la concepción existente sobre la gestión de las unidades y las características de la toma de decisiones de las escuelas. Se observa un cambio sustantivo que va desde la desconcentración hacia un incremento de la autonomía de las escuelas y de sus capacidades de decisión.

En cuanto a la naturaleza del proceso hemos relevado cuatro dimensiones claves: currículum, pedagogía, profesor y evaluación. Cada una de estas dimensiones son conceptualizadas y valoradas de diferente manera en los ciclos considerados. En el tema que nos preocupa en este artículo hay cambios sustantivos en los periodos definidos en cuanto a la concepción del conocimiento; objeto y metodologías de la evaluación.

La Tabla 1 resume las características de los movimientos de reformas en base a los criterios señalados.

Estos ciclos no se suceden en forma lineal y en cada país se viven de un modo particular. Cada uno de estos movimientos de reforma se caracteriza, a su vez, por preocupaciones particulares de evaluación y que son tratadas con el desarrollo teórico y metodológico propio de su tiempo. En los acápites siguientes compararemos los ciclos de reformas centrándonos en sus relaciones con la evaluación.

TABLA 1. CICLOS DE REFORMAS EN AMÉRICA LATINA

	Ciclos		
	Primera (80's)	Segunda (90's)	Tercera (2000)
Control	Recursos	Curriculum	Resultados
Actor clave	Privados/Municipios/Prov	Estado/Universidades	Escuelas/Ciudadanos
Política eje	Descentralización	Cambio curricular	Calidad/regulación
Procesos clave	Gestión	Interacción aula	Rendición de cuentas
Gestión Unidades	Desconcentradas	Descentralizadas	Autónomas
Decisiones Escuela	Dependientes	En consulta	Autónomas
Curriculum centrado en	Información	Conocimientos	Competencias
Pedagogía	Directiva	Directiva-Flexible	Flexible
Profesor	Funcionario/dependiente	Profesional/regulado	Profesional/Autónomo
Evaluación	Eficiencia	Aprendizajes absolutos	Aprendizajes relativos
Medición en base a	normas	criterios	Estándares

Fuente: Elaboración propia

1.1. Objetos y enfoques de la evaluación en los ciclos de reformas

La primera generación de reformas está centrada en la oferta del sistema y en la eficiencia de su funcionamiento. Es un periodo en el cual predomina un enfoque económico centrado en el estudio de costo y beneficios de las acciones emprendidas. Así, gran parte de las evaluaciones propias de esta primera generación, se concentran en la *eficiencia* del sistema y en donde las evaluaciones de aprendizaje se realizan, principalmente, en base a *normas* y que ordenan la población en función del promedio y porcentaje de respuestas correctas. Es un periodo, a la vez, de diseño y organización de sistemas nacionales de evaluación¹.

En la segunda generación de reformas parte importante de los estudios se concentra en los resultados de aprendizaje y en los factores que parecen explicarla. En ellas, las evaluaciones tienen por objeto los *aprendizajes absolutos* y las principales variables explicativas de las diferencias. Se inicia una evaluación de aprendizaje en base a *criterios* centrada en el logro de conocimientos, destrezas y habilidades. El foco está en la escuela y en los factores asociados que inciden en sus resultados. Se realizan estudios en torno a escuelas eficaces; observación y evaluación de prácticas pedagógicas, evaluación de gestión directiva, entre otros. En esa fase se consolidan los sistemas de evaluación nacionales; algunos países de la región participan en pruebas internacionales tales como TIMSS y PISA y se organiza el Laboratorio Latinoamericano de evaluación de la calidad de la educación (LLECE) patrocinado por la Oficina Regional de UNESCO. Desde el LLECE se organiza e implementa el Primer Estudio Comparativo Regional (1998), que da cuenta de los aprendizajes y desempeños en Lenguaje y Matemática de los estudiantes de América Latina y el Caribe (LLECE, 2001; Murillo y Román, 2008)

En la tercera generación de reformas los énfasis se encuentran en los resultados y en la rendición de cuentas. Esta es una época de grandes desafíos metodológicos para la investigación y los sistemas de uso de la información. La calidad de los resultados de aprendizajes sigue siendo foco principal de preocupación pero, al mismo tiempo, se destaca el estudio de los procesos pedagógicos y de aprendizajes en relación con los contextos sociales y culturales de los estudiantes.

¹ En el año 1991 sólo Costa Rica, México, Chile y Colombia contaban con sistemas nacionales de evaluación. Hoy día existen en la mayoría de los países (Wolf, L.2006).

Es un periodo en el cual adquieren importancia nuevos modelos teóricos y herramientas metodológicas y que, entre otros, permiten evaluar el progreso del aprendizaje de los alumnos y el "valor agregado" por la escuela. El *valor agregado* mide los resultados en función de los alumnos que la escuela atiende. Es decir, más que comparar promedios de un establecimiento con otro muy diferente se compara el progreso que tiene el alumno y el establecimiento con respecto a una medición inicial y su proyección (Froemel, 2003). La comparación es consigo mismo (Ravela, 2000). En Argentina, Brasil, Chile y Uruguay, entre otros, se observan avances importantes en esta dirección.

Por otra parte, esta nueva generación de reformas tiene un concepto más amplio de calidad educativa lo que plantea desafíos metodológicos en cuanto al objeto e instrumentos de evaluación. Se asume que los aprendizajes son cognitivos, valóricos y sociales; ocurren a lo largo de la vida y dependen de factores que son de naturaleza multivariada. Se destaca la necesidad de evaluar, entre otros, la calidad de los procesos en el aula; el desempeño de los profesores y el trabajo de los directivos (Murillo y Román, 2010).

2. CAMBIOS EN LOS PARADIGMAS DE LA EVALUACIÓN

Las generaciones de reformas coinciden, a su vez, con cambios en las teorías y desarrollo metodológico de la evaluación. Las teorías actuales superan los determinismos sociales, biológicos y culturales que caracterizaron gran parte de las evaluaciones de hace 20 o más años atrás. Hoy días se trabaja con modelos complejos y las explicaciones se buscan en los procesos internos de la escuela; en las prácticas pedagógicas y en las subjetividades de los actores (Shepard, 2000; Tylor *et al.*, 1996; Tedesco, 1998).

La mayor parte de las evaluaciones realizadas se inscriben en la tradición positivista; han medido el cumplimiento de objetivos y privilegian el diseño experimental. Estos estudios demostraron, a su vez, la importancia del contexto social como factor explicativo de los resultados de aprendizaje. En efecto, en gran parte de las evaluaciones realizadas en América Latina se constata que los factores sociales y el capital cultural de las familias es el principal factor explicativo de los problemas de aprendizaje². La escuela reproduce las desigualdades sociales originadas en la estructura social y permanecen en una "caja negra" los factores asociados a los procesos de aprendizaje que ocurren en la escuela (Tedesco, 1998; Buendía *et al.*, 2004; Ravela *et al.*, 2000).

Frente a este paradigma, se desarrolló en la Región una teoría diferente y que se consolidó hacia fines de los 80's.

En América Latina los trabajos de E. Schiefelbein y J. Farrel, (1975) demuestran que en países pobres la escuela *sí hace la diferencia* y tiene un mayor impacto que en las sociedades desarrolladas. En los sectores pobres la escuela primaria es más efectiva porque la situación de aprendizaje y las posibilidades de adquisición de saberes depende más de la escuela que de la socialización y cultura familiar (Schiefelbein, and. Farrel, 1973, 1975; Schiefelbein, 1984).

El cuestionamiento a los paradigmas de investigación estructuralista y el desarrollo de evaluaciones centradas en la escuela contribuyeron enormemente a poner más atención en los procesos internos que

² Esta conclusión se apoya en diversas investigaciones basadas en los estudios de Coleman (1966) realizados en los Estados Unidos; en los trabajos de Bourdieu y Passeron (1964) en Francia y en el desarrollo de una importante teoría estructuralista que, entre otros, fue promovida en la región por la misma CEPAL.

en las determinaciones externas a la interacción pedagógica. Gran parte de las reformas de los años 90 y que están centrados en los cambios “hacia adentro” del sistema, se inspiran en una hipótesis similar.

A fines de los años 90, este tipo de estudios recibe la influencia de los enfoques de las “escuelas efectivas” y de amplia trayectoria en los países desarrollados (Creemers, 1994; Slavin, 1996; Reynolds, 2000; Scheerens y Bosker, 1997).

Se entiende por escuela efectiva aquella que logra que sus alumnos alcancen metas educativas de calidad independientemente de su origen social. Estos estudios ponen el acento en los procesos y factores asociados al aprendizaje vinculados a la cultura escolar y a las prácticas pedagógicas.

Estudios realizados en América Latina demuestran que, una vez controlados los factores socio culturales asociados a la familia de los niños, los aprendizajes dependen centralmente de la calidad de la interacción y del proceso pedagógico que ocurre en las escuelas (Brunner *et al.*, 2006; Martinic, 2003; Murillo, 2003; 2005; 2007; UNICEF, 2005). Esta perspectiva plantea enormes desafíos a la investigación evaluativa. Entre ellos, exige construir modelos analíticos complejos y que consideren el contexto social y los procesos internos de las escuelas.

Los estudios fundantes de esta perspectiva identifican cinco grandes variables de la “caja negra” y que están relacionadas con mejores resultados de aprendizaje. Estas son: un liderazgo educativo fuerte, altas expectativas en cuanto a los resultados académicos de los alumnos, énfasis en las destrezas básicas, un clima seguro y disciplinado y evaluaciones frecuentes del progreso del alumno (Edmonds, 1979; Mortimore *et al.*, 1988; Reynolds *et al.*, 1994).

Sin embargo, estas variables no necesariamente son las mejores para explicar los procesos y resultados de la escuela en América Latina. Existe en la región un desafío de adaptación y de validación de nuevos modelos. Por ejemplo, estudios que han revisado la investigación sobre el tema demuestran que, en contextos de pobreza, los principales factores asociados a la calidad de los aprendizajes son: objetivos centrados en el aprendizaje de los alumnos; prácticas pedagógicas afectivas y efectivas; participación de los padres; apoyo al desarrollo profesional docente; trabajo directivo en equipo y altas expectativas de aprendizajes de los estudiantes (Muijs, D., 2003; Martinic, S. 2003; Raczynski D., G. Muñoz, 2005; Román, 2008; 2006; Sandoval, A. *et al.*, 2007; Zorrilla, M., *et al.*, 2007).

En síntesis, en las últimas tres décadas se observa una interesante evolución en los estudios evaluativos en la región y que deja atrás las explicaciones externas para estudiar con más importancia, los procesos internos; las interacciones y subjetividades de los actores en los contextos del establecimiento y de la sala de clases. En estas interacciones descansa, precisamente, la posibilidad del cambio y del mejoramiento en la escuela pese a las condicionantes que imponen las desiguales condiciones de vida de los niños.

3. CAMBIOS EN LAS METODOLOGÍAS E INSTRUMENTOS

Desde el punto de vista metodológico estos cambios se traducen en el diseño de modelos más complejos de análisis; se ha desarrollado la evaluación multiniveles y muchos estudios comienzan a trabajar con diseños mixtos integrando estrategias y técnicas tanto cuantitativas como cualitativas (Froemel, 2009).

Por otra parte, ha existido una interesante discusión sobre la calidad de las pruebas de evaluación y sus efectos en el sistema educativo. Entre las críticas destacan: las pruebas miden el aprendizaje como conocimiento adquirido y no como competencia o desempeño; no se evalúa la capacidad de raciocinio y

de resolución de problemas; no consideran las habilidades comunicativas orales y escritas, entre otras (Pini, M.E, 2000; Kelleghan, Th., 2001, Ravela, P., *et al.*, 2006). Por otra parte, se ha demostrado, que las pruebas nacionales de selección múltiple inciden en la simplificación del currículum ya que los profesores comienzan a enseñar las materias que serán medidas y entrenan a los estudiantes en ejercicios y rutinas de respuesta (Kelleghan and Greaney, 2001; Ravela, P, 2001; Darling-Hammond, L. 2004).

Por último, y esta es una de las consecuencias más complejas en la región, la difusión de los resultados de las pruebas sancionan a las escuelas si obtienen un bajo puntaje. Ello produce, una tendencia a seleccionar a los mejores alumnos promoviendo, además, el abandono o cambio de la escuela con el fin de que los puntajes promedio sean mejores.

Para superar estos problemas es importante un cambio en el tipo de constructo y desarrollar pruebas que permitan valorar los aprendizajes en función de los grados de dominio o nivel de logro alcanzado por los estudiantes (Ravitch, D., 1996; Baker, E.; Linn, R., 2002; Darling-Hammond, L. 2004).

Para enfrentar las consecuencias negativas del uso de pruebas que tienen altas consecuencias se recomienda el desarrollo de evaluaciones que den cuenta del progreso de los aprendizajes y controlar los factores de contexto asociados a la realidad social de los estudiantes. Este es el valor añadido o valor agregado y que calcula el desempeño de una escuela utilizando datos contextualizados y comparando resultados en distintos momentos de la trayectoria escolar. Se analiza así cuanto progresan los estudiantes comparándose al grupo con sus propios datos de ingreso o línea de base.

Además de los avances en las teorías de aprendizaje existe un amplio consenso en que la evaluación de la calidad de la educación no debe estar centrada, exclusivamente, en los aprendizajes cognitivos de los estudiantes. Para la UNESCO, por ejemplo, una educación de calidad no solo es "aprender a conocer" sino que también es "aprender a ser" y "aprender a convivir". Sin duda que una educación de calidad es la que logra una adecuada síntesis entre el aprendizaje de materias y el desarrollo de las áreas mencionadas. Este concepto plantea desafíos metodológicos en cuanto a la evaluación de otras áreas y dimensiones del desarrollo de los niños.

Por otra parte, se hace necesario evaluar no sólo los resultados sino que también los procesos educativos. En efecto, la calidad debe considerar el aprendizaje y otros factores que den cuenta de la efectividad de los procesos educativos desarrollados en los establecimientos y que conducen a mejores aprendizajes. La ampliación del foco y del objeto de la evaluación supone también la existencia de otros instrumentos que permitan, por ejemplo, evaluar el desempeño docente y directivo; las prácticas en la sala de clases; y la escuela como institución (Murillo y Román, 2010). Es recomendable, a su vez, que estos recojan información cuantitativa y cualitativa de modo de dar cuenta de la cultura escolar en toda su complejidad.

Las preocupaciones señaladas favorecen nuevas relaciones entre los métodos cuantitativos y cualitativos. Los métodos cuantitativos se aplican con mayor complejidad, se asume que las variables no actúan de un modo lineal y aditivo sino que con una fuerza sinérgica; como una red de interrelaciones entre factores en un contexto determinado. En los 90 este tipo de enfoques alcanza un mayor desarrollo gracias a la aplicación de nuevas técnicas estadísticas basada en los modelos jerárquicos lineales y en el diseño de modelos analíticos más amplios y comprensivos.

Por su parte la evaluación cualitativa también ha alcanzado un importante desarrollo en América Latina. Diversos estudios han demostrado la importancia de las representaciones; las interacciones y los

contextos culturales, entre otros, en los cuales se desenvuelven las prácticas pedagógicas; las políticas y programas de innovación educativa. Se ha constatado, por ejemplo, que la cultura de la escuela; las expectativas que tienen los profesores sobre sus alumnos y el clima interno, entre otras dimensiones simbólicas, constituyen importantes factores de calidad de los aprendizajes (Carlson, 2000; Martinic, 2003; 2005; Muijs, 2003; UNICEF, 2005; Román 2008; Murillo, 2007). El desarrollo de los enfoques etnográficos; de los análisis del discurso en el aula y el uso del video como registro han favorecido la comparación y el estudio cualitativo de los procesos de enseñanza y de las practicas en el aula en distintos contextos culturales.

4. USO DE RESULTADOS DE LA EVALUACIÓN

La puesta en marcha de los sistemas de evaluación tiene, como uno de sus supuestos, el hecho de que una vez conocidos los resultados, éstos serían ampliamente empleados por los usuarios en distintos ámbitos decisionales. La teoría de la acción subyacente es que la información produce cambios cognitivos y en la práctica de los sujetos. Desde esta perspectiva, la mayoría de los sistemas de evaluación se organizan con el objetivo de aportar información, al menos, a cuatro públicos específicos que la requieren para la toma de decisiones: autoridades; directivos; docentes y familias.

Tradicionalmente se ha definido como principal destinatario de las evaluaciones a las autoridades y alta burocracia de los Ministerios de Educación. Se asume que estos funcionarios son los que toman las decisiones –o aconsejan a quienes las toman- y, por tanto, deben tener la mejor información posible.

Este tipo de enfoque, a su vez, asume que las decisiones descansan sobre una base objetiva y racional de argumentación. Mientras mas información y evidencia empírica mejor fundamentada estará la decisión. La evaluación se define así como la base racional de la decisión y se concibe como una actividad neutral, objetiva y que está por encima de los intereses particulares en juego en cualquier decisión política. La mayor parte de los estudios evaluativos han respondido a esta necesidad. Sin embargo, el impacto de estos en las decisiones ha sido menor de lo esperado.

Las decisiones suelen tomarse antes de concluir las investigaciones y en estas también influyen las preferencias y presiones de los grupos de interés. Por otra parte, las decisiones o la fuerza del consejo no siempre están en la cúspide del sistema sino que también en la base y en el involucramiento de los grupos de interés de la sociedad civil en la definición y discusión de la agenda pública.

En otra escala del sistema, esta teoría sostiene que los datos de las evaluaciones son importantes para directivos y docentes para producir cambios en sus procesos de trabajo y calidad de resultados. Sin embargo, se ha demostrado que las evaluaciones son muy poco utilizadas por directores y profesores y tienen un bajo efecto en las prácticas. En Chile, por ejemplo, los resultados de las evaluaciones nacionales no son muy utilizadas por directores y profesores y tienen pocas consecuencias en los cambios en la enseñanza y en los aprendizajes.

Por último, por lo general, la información no se comunica en forma clara a los padres, para que comprendan los alcances e importancia de los resultados alcanzados. La comunicación de resultados supone una serie de habilidades superiores y que, por lo general, no están presentes en los padres. Entre ellas, habilidades de comprensión de lectura y de conceptos matemáticos claves (p.e. promedio, porcentaje); lectura de tablas y gráficos, entre otros (Ravela, 2006a; Martinic *et al.*, 2008).

En síntesis, pese a la abundante información y desarrollo de los sistemas de evaluación hay insuficiencias en la comunicación hacia los actores y faltan interpretaciones que les ayuden a comprender los problemas que verifican los datos reunidos. Los formatos de comunicación de los resultados no consideran, por lo general, los lenguajes y realidades de los distintos actores de las comunidades educativas.

5. CONCLUSIONES

Los ciclos de las reformas educativas han incidido en el concepto y status de la evaluación educativa en América Latina. Estos coinciden, a su vez, con cambios en las teorías y metodologías de la evaluación abriéndose un amplio campo de problemáticas y de estrategias para abordarlos.

En las últimas dos décadas hay cambios en las teorías y métodos de la evaluación educativa. De un paradigma positivista se transita a otro interpretativo. De una teoría del aprendizaje conductista a otra constructivista. De un énfasis en los insumos y factores externos se pasa a una preocupación central por los procesos escolares y las prácticas pedagógicas de la sala de clase. En los análisis de datos se ha extendido el uso de las técnicas de multinivel jerárquico.

Por otra parte, se ha demostrado la necesidad de mejorar la calidad y validez de los instrumentos y las tendencias se dirigen hacia la evaluación de los desempeños, estándares y de los procesos en el aula.

Las pruebas nacionales e internacionales han demostrado ser muy útiles para dar una visión del país o comparar un conjunto de países. Pero esta información no es suficiente para buscar soluciones y definir líneas de trabajo en escalas intermedias y locales de acción como lo requieren las provincias, los municipios y las escuelas. En estos niveles se necesita tener información sobre los factores que están asociados al logro de los aprendizajes y, a la vez, un modelo de análisis apropiado para la realidad local que ayude a definir prioridades y opciones para producir cambios en la gestión pedagógica de las escuelas.

Por ello es importante promover procesos descentralizados de evaluación y que aborden con la diversidad requerida los contextos y factores asociados a los logros. Este tipo de evaluación puede complementar las evaluaciones nacionales y aportar información que permita a los propios actores educativos tomar medidas que mejoren la educación en sus planteles.

Por otra parte, es importante promover estudios de carácter longitudinal de tipo panel. En estos estudios una misma muestra de alumnos es observada durante varios años con el objetivo de estudiar los cambios en su desempeño y los factores que inciden en ello.

Para enfrentar el problema de las comparaciones de alumnos con realidades muy desiguales es importante trabajar, como se ha señalado, desde una perspectiva de valor agregado. Ello se aplica con un amplio desarrollo en Inglaterra, Francia y en Estados Unidos, entre otros.

Esta información es de gran utilidad para reconocer en cada uno de los establecimientos o centros escolares la diversidad propia en los progresos de aprendizaje y, al mismo tiempo, conocer los principales factores en los cuales habría que intervenir para mejorarlo y que pueden variar de un establecimiento a otro o de acuerdo a perfiles y características de los propios estudiantes.

La evaluación no solo ha tenido cambios conceptuales y metodológicos sino que también un cambio en el lugar que ocupa en la sociedad. Desde una función asociada directamente a quienes diseñan y toman

decisiones se pasa a otra en la cual constituye una herramienta para actores diferentes en distintos momentos del diseño y ejecución de las políticas.

Las reformas educativas actuales viven un cambio de ciclo. Después de la reorganización institucional de los sistemas y del cumplimiento de la fase de cobertura el foco de los cambios está en los procesos educativos y en los factores que inciden en la calidad de los resultados.

Este cambio de ciclo también implica una transformación en las relaciones Estado-Sociedad y en la distribución del conocimiento sobre las políticas educativas. Las reformas educativas han dejado de ser pensadas y ejecutadas exclusivamente por el Estado y el concepto mismo de racionalidad, que subyace en las políticas, vive una profunda transformación. De una racionalidad técnica basada en el conocimiento experto y exclusivo del tomador de decisiones estamos pasando a una racionalidad que descansa en el diálogo, la argumentación y en la conversación.

Hoy día la agenda pública y las decisiones se tejen en espacios *conversacionales* más que en un gabinete o en la cúspide del poder³.

Lo público no es sólo la acción de la autoridad pública o a la presencia y acción Gubernamental y del Estado en la sociedad (Mény, Y y J-C Thoenig, 1989). Por el contrario, lo público comienza a entenderse como el encuentro de la racionalidad estatal con la voluntad social. Es el gobierno en y desde un contexto de pluralidad amplia. Lo público es una intersección, es un espacio de todos (Laville, J-L.2005; Pestoff, V., Osborne, S.P., et T. Brandsen, 2006; Vaillancourt, Y, 2007 ; Patton, M. 1997 ; Scriven. M., 1997).

En este tipo de proceso las actividades de seguimiento y evaluación son fundamentales para el desarrollo de los discursos de los actores y la construcción de las mejores soluciones y de políticas a los problemas que se deben enfrentar.

REFERENCIAS BIBLIOGRÁFICAS

- Arancibia, V. (2000). Factores que afectan el rendimiento escolar de los pobres. En E. Cohen. *Educación, conocimiento y equidad*. Santiago de Chile: Sur Ediciones.
- Artiles, A. (1994). La evaluación de los procesos de pensamiento de los maestros en contextos urbanos un estudio de caso en escuelas primarias de Guatemala. Ponencia presentada en la *Conferencia Anual de la Comparative and International Education Society*. San Diego, CA.
- Baker, E., Linn, R. (2002). *Validity issues for accountability systems*. Center for the Study of Evaluation. Los Angeles, CA. Disponible en: <http://www.cse.ucla.edu/CRESST/Reports/TR585.pdf>.
- Bellei, C, Muñoz, G., Raczynski, D., Pérez, L.M (2003). Escuelas Efectivas en Sectores de Pobreza: ¿Quién dijo que no se puede?. Hevia, R., *La Educación en Chile Hoy*, pp.347-372. Santiago de Chile: Ediciones Universidad Diego Portales.
- Bourdieu, P., Passeron, J.C. (1964). *La reproduction*. Paris: Ed. Minuit.
- Brunner, J.J. y Elacqua, G. (2006). Calidad de la educación. *Claves para el debate*. Santiago: Ril editores.

³ Conversar es "converger en algo", "estar juntos en algo".

- Cabrero, E. (2000). Usos y costumbres en la hechura de las políticas públicas en México. Límites de las policy sciences en contextos cultural y políticamente diferentes. *Revista Gestión y política pública*, 11(2), pp. 189-231 México: CIDE.
- Carlson, B. (2000). ¿Qué nos enseñan las escuelas sobre la educación de los niños pobres de Chile?. En *Revista de la CEPAL*, 72, pp.165-184.
- Coleman, J. et al., (1966). *Equality of educational opportunity*. Washington, D.C.: U.S. Office of Educational, National Center for Educational Statistics.
- Concha, C. (1996). *Escuelas efectivas en Chile: estudio de 32 escuelas exitosas en logros académicos y de alta vulnerabilidad*. Tesis de Magister en Educación, mención en Administración Educacional. Santiago de Chile: Pontificia Universidad Católica de Chile.
- Creemers, B.P.M. (1994). *The effective classroom*. London: Cassell.
- Edmonds, R. (1979). Effective School for the Urban Poor. *Educational Leadership*, 37, pp. 15-27.
- Froemel, J.E. (2009). La Efectividad y la Eficacia de las Mediciones Estandarizadas y de las Evaluaciones en Educación. *Revista Iberoamericana de Evaluación Educativa*, 2(1), pp. 10-28.
- Froemel, J.E. (2003). Evaluación de la calidad de la educación con equidad. El modelo de valor agregado. *Revista Persona y Sociedad*, 17(1), pp.165-178.
- Kelleghan, Th. and Greaney, V. (2001). The globalisation of assessment in the 20th century. *Assesment in education*, 8, pp. 87-102.
- Darling-Hammond, L. (2004). Standards, accountability, and school reform. *Teachers College Record*, 106(6), pp.1047-1085. Edición española parcial "Estándares, accountability y reforma escolar". Biblioteca virtual de Red Iberoamericana de Investigación sobre Cambio y Eficacia Escolar (Rinace). Disponible en: <http://www.rinace.org/>
- Darling-Hammond, L. (2005) Estándares, Accountability y Reforma Escolar., *Seminario Universidad Alberto Hurtado*. Escuela de Educación, Universidad de Stanford: PREAL.
- Hanson, M. (1997). *La descentralización educacional: problemas y desafíos*. Santiago de Chile: PREAL.
- Laville, J-L. (2005). *Sociologie des services. Entre marché et solidarité*. Paris: Eres.
- 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.
- Martinic, S. y M. Pardo (2003). *La investigación sobre eficacia escolar en Chile*. En: Murillo, J. (2003).
- Martinic, S. (2003). Representaciones de la desigualdad en la cultura escolar. *Revista Persona y Sociedad*, 17(1), pp. 129-146., Universidad Alberto Hurtado-Ilades.
- Martinic, S., Sebastián, C., Cortes, F., Salinas, V., Larrain T. (2008). Estrategia de comunicación y análisis de resultados SIMCE para la acción. *Camino al Bicentenario: propuestas para Chile*, pp. 19-46. Santiago: Vicerrectoría de Comunicaciones, PUC.
- Mény, Y. y Thoenig, J.C. (1989). *Les politiques publiques*. Paris: PUF.
- Mizala, A., Romaguera, P. (1998). Desempeño escolar y elección de colegios: la experiencia chilena., *Documentos de Trabajo, Serie Economía* (36). Santiago de Chile : Universidad de Chile.

- Muijs, D. (2003). La mejora y la eficacia de las escuelas en zonas desfavorecidas: resumen de resultados de investigación. *Revista electrónica Iberoamericana sobre Calidad, eficacia y cambio en Educación*, 1(2), Consultada el 10 de abril de 2003 en www.ice.deusto.es/rinace/reice.
- Murillo, F.J. (coord.) (2003). *La investigación sobre Eficacia Escolar en Iberoamérica. Revisión internacional del estado del arte*. Bogotá: Convenio Andrés Bello.
- Murillo, F.J. (coord.) (2004). La mejora de la escuela. *Cuadernos de Pedagogía*, 339, pp. 48-79.
- Murillo, J (2005). La Investigación en Eficacia Escolar y Mejora de la Escuela como motor para el incremento de la Calidad Educativa en Iberoamérica. *Revista Electrónica Iberoamericana sobre Calidad, Eficacia y Cambio en Educación*, 3(2).
- Murillo, F.J. (2007). *Investigación Iberoamericana sobre Eficacia Escolar*. Bogotá: Convenio Andrés Bello.
- Murillo, F.J. y Román, M. (2010). Retos en la evaluación de la calidad de la educación en América Latina. *Revista Iberoamericana de Educación*, 53, pp. 97-120.
- Murillo, F.J. y Román, M. (2008). [Resultados de aprendizaje en América Latina a partir de las evaluaciones nacionales](#). *Revista Iberoamericana de Evaluación Educativa*, 1(1), pp.6-35.
- Ozlak, O. (1997). Estado y Sociedad. Colección CEA-CBC, *Las nuevas reglas de juego*, 1. Buenos Aires: UBA.
- Ozlak, O. (1999) Quemar las naves (o cómo lograr reformas estatales irreversibles). Ponencia presentada al *IV Congreso Internacional del CLAD*, 19-22 de Octubre 1999. México.
- Patton Quinn, M. (1997). Toward distinguishing empowerment evaluation and placing it in a larger context. *Evaluation Practice*, 18(2), pp. 147-163.
- Pini, E.M., (2000). Lineamientos de política educativa en los Estados Unidos: debates actuales. Significados para A.Latina. *Education Policy Analysis Archives*, 8(18). Disponible en: <http://epaa.asu.edu/epaa/v8n18.html>.
- Pestoff, V., Osborne, S.P., et Brandsen, T. (2006). Patterns of co-production in public services. Some concluding thoughts. *Public Management Review*, 8(4), pp. 591-595.
- PREAL (2006). *Sobre estándares y evaluaciones en América Latina*. Grupo de trabajo sobre estándares en Evaluación. Santiago: PREAL.
- Ravela, P. (2006a) ¿Cómo presentan sus resultados los sistemas nacionales de Evaluación Educativa en América Latina? *PREAL*, 22, pp. 181-248. Santiago de Chile: PREAL.
- Ravela, P., Wolf, L., Valverde, R., Esquivel, J.M. (2006b). *Los próximos pasos: ¿Cómo avanzar en la evaluación de aprendizajes en América Latina?*, pp.53-122. Santiago de Chile: PREAL.
- Ravitch, D. (1996). *Estándares nacionales en educación*. Santiago de Chile: PREAL.
- Raczynski D., Muñoz, G. (2005). *Efectividad escolar y cambio educativo en condiciones de pobreza en Chile*. Santiago de Chile: Mineduc.
- Reimers, F. (2001). *Unequal Schools, Unequal Chances*. Cambridge, MA: Harvard University Press.
- Reimers, F. (2002). Tres paradojas educativas en América latina. Sobre la necesidad de ideas públicas para impulsar las oportunidades educativas. Documento preparado para el *Diálogo regional en Educación Junio 2002*. BID.

- Reynolds, D. (1989). School Effectiveness and School Improvement. A Review of the British Literature. D. Reynolds; B. Creemens (eds). *School Effectiveness and Improvement*. London: RION.
- Román, M. (2008). Investigación Latinoamericana sobre Enseñanza Eficaz, ILEE. En UNESCO, *Eficacia escolar y factores asociados en América Latina y el Caribe*, pp.209-225. Santiago de Chile: UNESCO.
- Román, M. (2006). El desafío de aprender a leer, comprender y razonar en escuelas vulnerables. *Pensamiento Educativo*, 39, pp.69-86.
- Sandoval, A. y Barrón, J.C. (2007). El programa de investigación del movimiento de escuelas eficaces: hacia una perspectiva basada en los actores en el contexto de América Latina. *Revista Electrónica Iberoamericana sobre Calidad, Eficacia y Cambio en Educación*, 5(5e), pp. 264-270. Consultado el 10 mayo 2008 en en: <http://www.rinace.net/arts/vol5num5e/art36.pdf>.
- Scriven, M. (2003). Evaluation in the new millennium: the transdisciplinarity vision. Donaldson, S.I. and Scriven, M.(eds) *Evaluating social programs and problems: vision or the new millennium*. New Jersey: Lawrence Erlbaum.
- Scheerens, J. (1992). *Effective schooling:research, theory and practice*. Londres: Casell.
- Schiefelbein, E. (1984). La investigación sobre la calidad de la enseñanza en América Latina. En Revista *La Educación de OEA*, 28(96), pp. 88-116.
- Schiefelbein, E., Farrel, J. (1982). *Eight Years of Their Lives: Trough Schooling to the Labor Market in Chile*. Ottawa: IDRC.
- Schiefelbein, E., Farrel, J. (1973). Expanding the Scope of Educational Planning: the Experience of Chile. *Interchange*, 2(1), pp. 18-30.
- Schleicher, A. (2006). La mejora de la calidad y de la equidad en la educación: retos y respuestas políticas. Polanco, J. (ed). *Políticas educativas de éxito: análisis a partir de los informes PISA*, pp.11-47. Madrid: Santillana.
- Shepard, L., (2000). The role of assessment in a learning cultur. *Educational Researcher*, 29(7), pp. 4-14.
- Solari, A. (1994). La desigualdad educativa: problemas y políticas. *Serie Políticas Sociales* (4). Santiago, CEPAL.
- Tedesco, J.C., (1998), Desafíos de las Reformas Educativas en América Latina. *Revista Paraguaya de Sociología*, 35(2). pp. 7-18.
- Teddlie, C. and Reynolds, D. (2000). *The International Handbook of School Effectiveness Research*. London and New York: Falmer Press.
- Tylor, C. y Lyons, L. (1996) Theory-based Evaluation. *Evaluation practices*, 17(2), pp 177-184.
- UNICEF (2004). *Escuelas Efectivas en sectores de pobreza: ¿Quién dijo que no se puede?*.
- Unesco (2000). Laboratorio Latinoamericano de Evaluación de la Calidad de la Educación. *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.
- Vaillancourt, Y., (2007). *La co-construction des politiques publiques avec l'apport de l'économie sociale*. (juillet 2007 m.s)
- Wolf, L. (2006), Las evaluaciones educacionales en América Latina. Avance actual y futuros desafíos. *PREAL* (2006), pp. 13-52.

Zorrilla, M. y Ruíz, G. (2007). Validación de un modelo de Mejora de la Eficacia Escolar en Iberoamérica. Factores de la escuela: cultura para la mejora. El caso de México. *Revista Electrónica Iberoamericana sobre Calidad, Eficacia y Cambio en Educación*, 5(5e), pp. 200-204. <http://www.rinace.net/arts/vol5num5e/art26.pdf>.



**LA NUEVA NATURALEZA DEL APRENDIZAJE Y DE LA
EVALUACIÓN EN EL CONTEXTO DEL DESARROLLO
COMPETENCIAL, RETOS EUROPEOS EN LA EDUCACIÓN
DEL SIGLO XXI**

**THE NEW NATURE OF LEARNING AND OF EVALUATION IN THE CONTEXT OF
COMPETENTIAL DEVELOPMENT, EUROPEAN CHALLENGES IN THE
EDUCATION OF THE 21ST CENTURY**

Joan Mateo y Dimitris Vlachopoulos

Revista Iberoamericana de Evaluación Educativa 2010 - Volumen 3, Número 3

<http://www.rinace.net/riee/numeros/vol3-num3/art3.pdf>

Fecha de recepción: 05 octubre de 2010
Fecha de dictaminación: 16 octubre de 2010
Fecha de aceptación: 20 octubre de 2010

«Los profesores seremos felices cuando renunciemos a enseñar e intentemos simplemente dar a nuestros alumnos marcos de aprendizaje que les permitan reflexionar, hacer trabajar los ojos, las orejas, el cerebro y por qué no el cuerpo, la creatividad. La crisis actual de la enseñanza se da porque renunciemos a entrar en esta nueva lógica (François Bégedeau, "El Periódico" 6 de enero de 2009, p.49).

En 1991 y por primera vez en Estados Unidos, el gasto total en bienes industriales fue superado por el aplicado a las tecnologías de la información y la comunicación (Trilling, B. y Fadel, Ch. 2009). Con ello se materializa un cambio similar al que se produjo a mediados del siglo pasado cuando se pasó de la sociedad agrícola a la industrial, en este caso la deriva se produce desde la sociedad industrial a la del conocimiento.

En este contexto cabe pensar que los sistemas educativos deben prepararse para generar respuestas nuevas acordes con las nuevas necesidades. El conocimiento actual, el del siglo XXI, mantiene básicamente su esencia clásica pero presenta algunas características nuevas que le otorgan una naturaleza distinta al de épocas anteriores.

La construcción del saber postmoderno se fundamenta en el análisis crítico de una ingente acumulación de datos e información que circula por la red y que genera cambios sustanciales y vertiginosos del propio conocimiento. La magnitud de la tarea obliga a los investigadores a especializarse y a fragmentar los campos del saber para poder así atender su rápida obsolescencia y su continua y acelerada innovación.

Consecuentemente si la naturaleza del conocimiento se modifica, también debe hacerlo nuestra relación con él. A una relación de dominio, propia de la educación del siglo pasado le sucede una de gestión competente que marca los enfoques pedagógicos del actual.

La educación como vínculo entre el hombre y el universo a través del establecimiento de enlaces entre el conocimiento y la interpretación de la realidad para comprenderla e intervenir en ella, precisa de nuevas conceptualizaciones. Dice Borghesi : "educar es convertir el conocimiento en experiencia y la información en vida". (La Vanguardia 16 de Diciembre de 2005 –la contra- y en su libro "El sujeto ausente" (2005))

Entendemos que el filósofo y pedagogo italiano no hace sino reclamar la necesidad de visiones más integrales del universo y establece como camino para conseguirlo un tipo específico de nexo entre el conocimiento y la realidad, el marcado por la preeminencia de la experiencia y la vida respecto del conocimiento y la información. Creemos que en el análisis de las características de ese nexo reside la naturaleza de los nuevos enfoques formativos basados en el desarrollo competencial y la base de las nuevas concepciones pedagógicas y a ello dedicaremos este escrito.

1. EL CONCEPTO DE COMPETENCIA Y SU CARACTERIZACIÓN

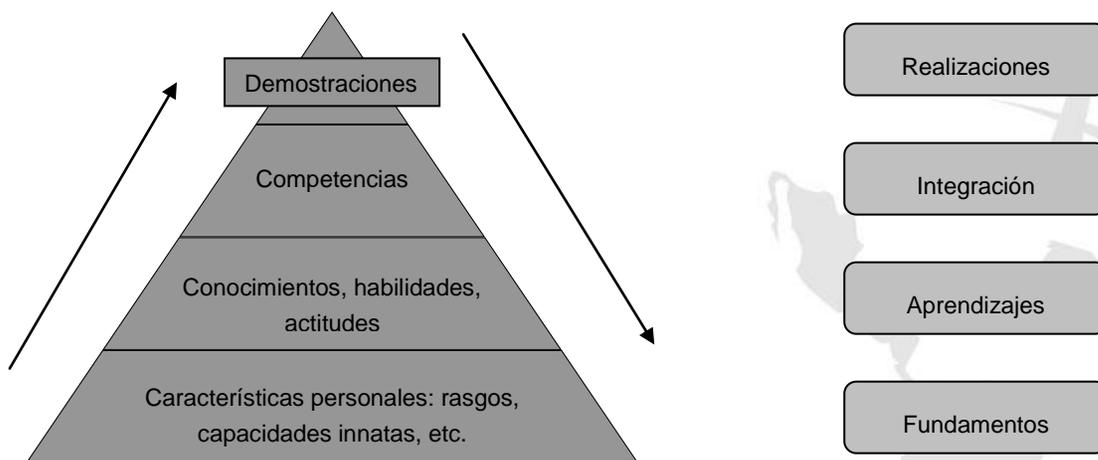
Para una mejor comprensión de su concepto, consideramos interesante revisar, cómo entendemos las competencias desde la vertiente educativa y cómo se estructuran según su función, ya que, revisando la literatura, observamos que las competencias se perciben desde diversos- y dispares- puntos de vista.

Más concretamente, en la bibliografía se han utilizado numerosos términos para describir los resultados de los procesos de aprendizaje, como habilidades, destrezas, capacidades, competencias, etc., sea como sinónimos o como términos con matices diferentes. De esta manera podemos afirmar: (Mateo, J., Escofet, A, Martinez, F y Ventura, J., 2009):

- Los rasgos y las características personales son los cimientos del aprendizaje, la base innata desde la que se pueden construir las experiencias subsiguientes. Las diferencias en rasgos y características ayudan a explicar por qué las personas escogen diferentes experiencias de aprendizaje y adquieren diferentes niveles y tipologías de conocimientos y habilidades.
- Los conocimientos, las habilidades y las actitudes se desarrollan a partir de las experiencias de aprendizaje, que, si se definen de una manera amplia, incluyen tanto la escuela como el trabajo, la familia, la participación social, etc.
- Las competencias son combinaciones de conocimientos, habilidades y actitudes adquiridas. Se desarrollan a partir de experiencias de aprendizaje integrativas en las que los conocimientos y las habilidades interactúan con el fin de dar una respuesta eficiente en la tarea que se ejecuta.
- Las demostraciones comportan la aplicación de las competencias aprendidas, en contextos específicos.

La siguiente figura muestra la estructura jerárquica de estos conceptos y permite establecer las diferencias¹.

FIGURA 1. LA ESTRUCTURA JERÁRQUICA DE LOS CONCEPTOS



¹ Esta conceptualización procede del trabajo realizado por el Council of the National Postsecondary Education Cooperative (NPEC) y su Working Group on competency-based Initiatives, patrocinado por el NCES (National Center for Education Statistics). Referencia: NCES (2002). *Defining and Assessing Learning: Exploring Competency-Based Initiatives*. Disponible en www.inces.ed.gov/publicsearch/ [consultado el día 3 de septiembre de 2008].

Presentamos para su caracterización un conjunto de diferentes definiciones que aparecen frecuentemente en la bibliografía sobre el concepto de competencias:

- “Competencia es la capacidad de actuar de manera eficaz en un tipo definido de situación, una capacidad que se sustenta en conocimientos, pero no queda reducida a éstos” (Perrenoud, 1999).
- “Competencia es un saber hacer complejo, resultado de la integración, la movilización y la adecuación de capacidades (conocimientos, actitudes y habilidades) utilizados eficazmente en situaciones que tengan un carácter común” (Lasnier, 2000).
- “La competencia es la habilidad aprendida para llevar a cabo una tarea, un deber o un rol adecuadamente. Un alto nivel de competencia es un pre-requisito de buena ejecución. Tiene dos elementos distintivos: está relacionada con el trabajo específico en un contexto particular, e integra diferentes tipos de conocimientos, habilidades y actitudes. Hay que distinguir las competencias de los rasgos de personalidad, que son características más estables del individuo. Se adquieren mediante el *learning-by-doing* y, a diferencia de los conocimientos, las habilidades y las actitudes, no se pueden evaluar independientemente” (Roe, 2002).
- “Competencia es más que conocimiento y habilidad. Implica la habilidad de responder a demandas complejas mediante la movilización de recursos psicosociales (incluyendo habilidades y actitudes). (Rychen, D.S. y Salganik, L.H. 2005)
- “Competencia es la gestión global y eficiente de la aplicación de un conjunto seleccionado e integrado de conocimientos y habilidades sobre contextos de realidad complejos con el objeto de interpretarlos, comprenderlos e intervenir eficazmente en ellos” (Mateo, J. y Martínez, F., 2008).

Sin embargo y tal como señala Weinert (2001), adquirir una competencia es algo más que poseer una habilidad respecto de un dominio. Implica regulación, monitorización y capacidad de iniciativa en el uso y desarrollo de dicha habilidad.

Desarrollar el uso competencial de un conocimiento o habilidad sobrepasa su mero dominio o la aplicación mecánica y habilidosa del mismo sobre un contexto de realidad. Supone interacción, mediación y gestión entre el conocimiento y la realidad física social y cultural y actuar con efectividad y eficiencia no tan sólo en la realización de la aplicación sino también en la interpretación del contexto y sus significados.

Hacer un uso competente del conocimiento o de una habilidad incluye saber lo que uno necesita saber y también lo que no se necesita. Llegar a este tipo de comprensión supone reconocer el contexto donde una habilidad funciona con efectividad en interacción con él.

Poseer una habilidad específica supone simplemente dotarnos de un instrumento para actuar en un contexto específico y en cambio la adquisición de una nueva competencia implica un efecto directo sobre nuestra capacidad para interpretar y dotar de nuevos sentidos a la realidad sobre la que actuamos.

En definitiva la adquisición de conocimientos y habilidades desde esta nueva dimensión competencial nos aporta nuevas metáforas que actúan de mediadoras en nuestras interacciones con la realidad y la dotan de nuevos y más relevantes significados.

A modo de síntesis, podemos decir que:

- La competencia es la combinación de habilidades, actitudes y conocimientos necesarios para desarrollar una tarea de manera eficaz.
- Las competencias se demuestran en la acción y, por lo tanto, sólo son evaluables en tanto que hay actividades que impliquen que se llevan a cabo.
- Las competencias son aprendidas y se desarrollan a partir de actividades que permiten integrar habilidades, actitudes y conocimientos aprendido anteriormente quizás de manera separada.

2. LAS COMPETENCIAS BÁSICAS EN EL CONTEXTO DE LA UNIÓN EUROPEA

El desarrollo de las competencias básicas forma parte esencial del programa de trabajo de la U.E. conocido como Education and Training 2010 y que surge como respuesta a los planteamientos del Consejo Europeo que se reunió en Lisboa el 23 y 24 de Marzo de 2000 y que posteriormente fueron ratificados en la resolución del año 2002 donde se establece definitivamente que las competencias básicas (key competences) constituirían una clara prioridad educativa para todas las edades y que se deberán desarrollar programas para su concreción e incorporación en los currícula de todos los países de la Unión.

Para dar respuesta a esta demanda se desarrollaron iniciativas para tratar de determinar cuáles eran las competencias básicas fundamentales que debía desarrollar cualquier estudiante durante la época de formación reglada a fin de garantizar su capacidad para insertarse como ciudadano en la sociedad. Se pretendía con ello orientar los currícula alrededor de ellas.

Este planteamiento condujo, inicialmente, a la generación de múltiples plataformas constituidas casi todas ellas por un gran número de competencias de carácter atomizado. Esta forma de actuar fue duramente criticadas (Perrenoud) pues suponían nuevamente situar el proceso educativo bajo un prisma conductista y retrotraían el sistema a los modelos orientados a objetivos, más propios de los años setenta. Bajo este enfoque las competencias eran definidas como la capacidad para realizar tareas específicas de modo individual (Eurídice, 2002)). Estos modelos únicamente han tenido un cierto desarrollo en el ámbito de la formación profesional manteniéndose una tradición de trabajo por objetivos específicos ahora ampliados a las competencias atomizadas. (Rodríguez, *et al.*, 2010)

Se dieron también enfoques de carácter cognitivista en las que las competencias son vistas como conjuntos de procesos mentales desarrollados por el estudiante con la finalidad de tener un desempeño efectivo. El individuo, bajo este paradigma es considerado como un sujeto activo procesador de información que posee competencias cognitivas que deben ser desarrolladas usando nuevos aprendizajes y habilidades estratégicas. Esta posición también fue objeto de crítica ya que los planteamientos curriculares que surgen de ellas pecan de incurrir en ignorar los contextos de aplicación.

Finalmente se ha ido abriendo camino las concepciones integradas y holísticas de las competencias que son las reflejadas en las definiciones escogidas por nosotros anteriormente.

Bajo esta acepción uno de sus principales fundamentos reside en el establecimiento de fuertes nexos entre competencia, individuo, tarea y contexto. Pero a la vez esta concepción insiste en hacer una distinción entre el conocimiento y el desempeño.

Como resultado de estas iniciativas, en el ámbito europeo se huye de la atomización y se definen para la enseñanza reglada ocho dominios competenciales de carácter muy amplio que se consideran como básicos para que una persona pueda lograr su realización personal, integrarse en la sociedad, ejercer la ciudadanía activa y ser capaz de desarrollar un aprendizaje permanente a lo largo de toda la vida.

Estos dominios son concebidos como grandes orientaciones del currículo y de los procesos de formación continua y deberán revisarse y actualizarse de forma periódica, pero siempre desde el espacio definido por los contenidos curriculares (conceptos, procesos, actitudes, valores) y el contexto.

Los ocho grandes dominios a los que hacemos referencias son:

- 1.- *Competencia en comunicación lingüística* entendida como habilidad para expresarse e interpretar pensamientos, sentimientos y hechos de forma oral y escrita e interactuar lingüísticamente de forma apropiada en cualquier contexto social y cultural.
- 2.- *Competencia en lenguas extranjeras* se entiende como una ampliación de la anterior y se desarrolla de acuerdo a las necesidades del contexto. También se incluye en esta competencia el desarrollo de habilidades orientadas a la mediación y a la comprensión intercultural.
- 3.- *Competencias en la matemática en la ciencia y en la tecnología* implican habilidades en el uso del cálculo matemático para la resolución de problemas cotidianos, en el uso del conocimiento y la metodología científica para explicar el mundo y la naturaleza y en la comprensión y aplicación de la tecnología para modificar los contextos naturales de acuerdo a las necesidades humanas.
- 4.- *Competencia digital i para el tratamiento de la información* implica el uso crítico de los medios electrónicos para el trabajo, la diversión y la comunicación. En su nivel más elemental supone usar la tecnología multimedia para obtener, evaluar, almacenar, producir, presentar e intercambiar información y para comunicarse y participar en las redes vía internet.
- 5.- *Competencia para aprender a aprender* implica la disposición y habilidad para organizar y regular el propio aprendizaje tanto de forma individual como colectiva. Incluye la habilidad para gestionar con eficiencia el uso del propio tiempo y para aplicar el conocimiento y las habilidades sobre una variedad sustantiva de contextos.
- 6.- *Competencias cívicas e interpersonales* abarcan todo tipo de conductas que se deben desarrollar a fin de ser capaz de participar de manera constructiva y eficiente en la vida social y en la resolución de conflictos tanto de carácter personal como social.
- 7.- *Competencias asociadas a la autonomía y a la iniciativa personal* constituyen componentes de carácter pasivo y activo e implican desarrollar la capacidad de aceptar responsabilidades para la propia acción, desarrollar visión estratégica, ser capaz de marcarse objetivos y conseguirlos. Induce en general a la aceptación activa de la innovación y el cambio.
- 8.- *Competencia para la expresión cultural* implica desarrollar la capacidad para conocer, comprender, apreciar y valorar críticamente diferentes manifestaciones culturales y artísticas, utilizarlas como fuente de enriquecimiento y disfrute y considerarlas como parte del patrimonio de los pueblos la importancia de la expresión creativa tanto de ideas como de experiencias o emociones.

Estos dominios competenciales se generaron asociados no tan sólo a la formación reglada sino también a los procesos de formación a lo largo de toda la vida y se deben percibir como interdependientes entre

ellos y en cualquier caso el énfasis se ha de situar en el desarrollo del pensamiento crítico, la creatividad, la iniciativa, la resolución de problemas, la valoración de los riesgos, la toma de decisiones y la gestión constructiva de los sentimientos y las emociones.

Constituyen el marco de referencia clave y de gran valor para el desarrollo de las competencias a implementar en los currícula de la enseñanza obligatoria de los países europeos. Tanto es así que todos los países de la U.E. han ido innovando sus propuestas curriculares clásicas, incorporando fundamentalmente conjuntos de competencias claramente inspirados o derivados de la propuesta marco.

En el caso concreto de España la propuesta se articuló a partir de las siguientes ocho dominios competenciales:

- Competencia en comunicación lingüística.
- Competencia matemática.
- Competencia en el conocimiento y la interacción con el mundo físico.
- Tratamiento de la información y competencia digital.
- Competencia social y ciudadana.
- Competencia cultural y artística
- Competencia para aprender a aprender.
- Autonomía e iniciativa personal

Dada la estructura autonómica de nuestro país cada autonomía, en el legítimo uso de sus derechos, ha optado por desarrollar su propia propuesta (inspirada, eso sí, en el marco de referencia europeo) o aceptar la estatal y, en algún caso, añadirle alguna competencia más de carácter propio.

3. COMPETENCIAS Y CURRÍCULO

La mayoría de países europeos han ido incorporando en los últimos años el desarrollo competencial en sus currícula. En España la Ley Orgánica 2/2006 de 3 de mayo, de Educación incorpora las competencias básicas al currículo, como uno de sus elementos, y les otorga el "rol" de referente curricular para la evaluación general de diagnóstico y la evaluación de diagnóstico.

Se asume con esta decisión legislativa, las conclusiones de los organismos internacionales (la OCDE y la Comisión Europea) y del programa para la evaluación internacional de los alumnos de la propia OCDE (PISA) sobre la enseñanza y aprendizaje de las competencias básicas como un medio para mejorar la calidad y la equidad de los sistemas educativos.

Sin embargo basta revisar la mayoría de las propuestas curriculares surgidas a partir de la incorporación de las competencias básicas para detectar que no se ha resuelto su integración en el conjunto del diseño curricular. Se trata más bien de una simple yuxtaposición en las que las competencias aparecen listadas sin ningún planteamiento de cómo encajarlas en el sistema.

Posiblemente la dificultad se centra en la conceptualización que se hace de las competencias considerándolas más como objetivos a asumir que potencialidades de los sujetos a desarrollar.

En el paradigma conductista los objetivos constituían los elementos clave en el modelo de transmisión-recepción y eran enunciados y prescritos de manera invariable mientras que en una educación basada en competencias estas deben ser establecidas y consideradas de acuerdo a las circunstancias del entorno y los elementos culturales en los cuales se encuentra inmerso el sujeto que aprende. (González, O., 2002)

Es justamente ese carácter contextual de las competencias el que les confiere su propia esencia y nos obliga a modificar todas nuestras acepciones pedagógicas. No podemos desarrollar un currículo y unos modelos de evaluación orientados exclusivamente a la consecución de unos objetivos de carácter invariante sino a desarrollar nuestra capacidad para leer, analizar e interpretar realidades distintas y variables y desde su comprensión saber gestionar eficientemente nuestros conocimientos por tal de intervenir eficazmente en ellas.

Con ello no negamos en absoluto la importancia del conocimiento y su aprendizaje para desarrollar nuestras competencias, sin conocimiento no hay ninguna posibilidad de desarrollo competencial, constituye una de sus condiciones necesaria pero no suficiente. Veamos cuál es el papel asignado a cada uno de los elementos básicos que intervienen en el desarrollo de las competencias.

El conocimiento es quien nos ilumina la realidad y nos ayuda a profundizar en su comprensión, las habilidades son las que favorecen la interacción entre conocimiento y realidad, las actitudes son las que dinamizan la interacción y los valores los que orientan a nuestra acción para que cobre significado. (personal, social, ético, etc.)

La clave de todo ello está en cómo se deben incorporar las competencias sugeridas por la UE o incorporadas definitivamente por nuestro país o comunidad en el currículo y de qué manera influirán en su diseño. A nuestro juicio una inclusión integrada de las competencias básicas en el currículo exige algunos cambios de enfoque.

En primer lugar se deben integrar los diferentes saberes, tanto los formales, incorporados a las diferentes áreas o materias, como los no formales. En segundo lugar, generar contextos de oportunidad para permitir a todos los estudiantes, desde esa asunción holística del conocimiento poder ponerlo en relación con ámbitos de la realidad diversos desde distintos tipos de situaciones y contextos y en tercero plantear la evaluación, en todos sus niveles, desde un nuevo prisma que la sitúa como una parte fundamental del propio proceso de aprendizaje orientado al desarrollo competencial.

Procederemos brevemente a reflexionar respecto de las dos primeras exigencias dentro de este apartado para analizar, en mayor profundidad, el nuevo papel de la evaluación en el siguiente.

Para trasladar al currículo la necesidad de integración de los saberes debemos concebirlo de una manera mucho más dinámica del que aparece en la mayoría de diseños europeos. Concretamente entendemos que la propuesta curricular no puede limitarse a unos listados de contenidos clasificados por materias.

Se debe abrir un diálogo de carácter intra e interdisciplinar y analizar las relaciones más importantes existentes entre los contenidos dentro de una propia disciplina y entre las distintas disciplinas. El currículo debería explicitar estas relaciones estableciendo verdaderos mapas conceptuales entre las diferentes materias.

Esta es una tarea al alcance de los expertos que diseñan los currícula y que facilitaría enormemente la tarea del profesor en el aula cuando trata de organizar el aprendizaje de los distintos elementos y sobre todo para incrementar su capacidad para generar actividades que impliquen el uso integrado de ellos.

Por otro lado la sustitución de las opciones de competencias atomizadas de principio del 2000 por los ocho grandes dominios actuales, que consideramos acertada, reduce, sin embargo, el grado de operatividad para proceder a su desarrollo. Para solventar este escollo alabamos soluciones como la intentada por países como Finlandia en las que, en su propuesta curricular, asocian a cada uno de los grandes dominios las habilidades que se consideran más directamente relacionadas con la competencia a desarrollar y se establecen distintos niveles de profundización.

A modo de ejemplo hemos escogido del modelo finlandés un dominio, el denominado “Planificar y organizar actividades” a esta megacompetencia le asocian las siguientes habilidades:

- Planificar y organizar las propias actividades de trabajo.
- Hacer un correcto uso del tiempo y los recursos disponibles.
- Priorizar-
- Realizaciones monitorizadas.

Para las que establecen tres niveles de ejecución:

- Nivel 1: Ejecutar bajo supervisión.
- Nivel 2: Ejecutar con orientación.
- Nivel 3: Ejecutar de forma independiente iniciando y evaluando actividades complejas.

Sin entrar a juzgar la bondad de esta propuesta concreta, si que percibimos como muy positivo esta disección de cada uno de los dominios competenciales sugiriendo las habilidades implicadas en la competencia señalando los niveles de ejecución.

La segunda condición que debería contemplar el currículo sería prever tiempos y recursos docentes para el desarrollo de actividades complejas (trabajos interdisciplinares de síntesis, proyectos, etc.) que permitan poner en juego los dominios competenciales a niveles de ejecución realmente intensos.

Es evidente que el desarrollo competencial nos va a exigir cambios sustanciales en la programación. Posiblemente no podremos mantener los modelos basados en el temario en los que el tema constituye su unidad fundamental y éste presumiblemente será substituido por modelos basados en la programación por actividades.

Una actividad no es sino un contexto de oportunidad de aprendizaje. En una simulación de realidad con un cierto nivel de complejidad donde podemos intervenir desde la aplicación integrada de diversos conocimientos. Su realización exige activar las actitudes de la persona y permite trabajar colectivamente y desarrollar así un conjunto de habilidades que no encuentran fácil acomodo en otro tipo de contextos.

Al margen de las actividades diseñadas en el contexto del desarrollo de las diferentes materias, es conveniente tener presente espacios concebidos para poder trabajar de forma cooperativa y sobre realidades realmente complejas.

Finalmente señalamos que el trabajo en las áreas del currículo para contribuir al desarrollo competencial debe complementarse, al margen de lo expuesto, con otras medidas de carácter cultural, organizativo y funcional.

Así la organización de los centros y las aulas deberán acomodarse al nuevo paradigma flexibilizando sus normas incorporando medidas funcionales que permitan el trabajo individual y en grupo.

Se deberán incorporar nuevas metodologías y recursos didácticos, replantearse el papel de las TIC para contribuir a las competencias asociadas a la comunicación, el uso de los servicios, etc. Establecer modelos de acción tutorial que puedan contribuir a la regulación de los aprendizajes y al desarrollo emocional y social de los estudiantes.

Los centros deberían potenciar su trabajo en red con otros centros a fin de intercambiar experiencias y compartir recursos de todo tipo. Deberían también interaccionar con el entorno para integrar todas sus potencialidades didácticas convertirse en definitiva en verdaderos ecosistemas de aprendizaje.

Evidentemente estas medidas que superan la voluntad o campo de acción del profesorado son imprescindibles para que una docencia diferente pueda darse. Así pues, requieren en definitiva cambios en las políticas y estructuras de las instituciones (Cano, 2008).

A nivel docente, en tanto que la atención se focaliza en el aprendizaje más que en la enseñanza y en el rol activo por parte del estudiante para construir su conocimiento, esto nos obliga a un cambio cultural que pasa por la revisión del modo en que entendemos la función docente y la transmisión de contenidos y replantear el "qué" tienen que aprender los estudiantes y el "cómo" es la mejor manera de aprenderlo. Por supuesto, esta situación no es nueva. La mayoría de los docentes venían trabajando en esta línea por lo que se ha de partir de rescatar aquello que ya hacían y que les sirve dándole una dimensión colectiva.

4. LA EVALUACIÓN DE LOS APRENDIZAJES EN UN CONTEXTO DE CAMBIO

Tal como hemos visto la nueva educación orientada al desarrollo competencial de los estudiantes implica modificar profundamente nuestro pensamiento sobre formación, instrucción y docencia .pero también supone modificarlo de forma casi total den los planteamientos evaluadores,

Así en los últimos diez años se ha producido un conjunto muy importante de cambios en la misma naturaleza de la evaluación de los aprendizajes que están afectando al pensamiento actual con respecto al binomio enseñanza-aprendizaje y el papel de la evaluación. Dividimos dichos cambios en las tres categorías principales que se explican a continuación.

4.1. Cambios en el enfoque del proceso de enseñanza-aprendizaje

Se otorga más importancia a los procesos de aprendizaje que a los de enseñanza significa que el centro de gravedad se sitúa en los outputs más que en los inputs. Se cumple con eso uno de los principios básicos del nuevo paradigma organizativo de la educación, el de la primacía de las finalidades (Hutmacher, 1999), según el cual la acción se orienta de manera prioritaria a la consecución de los objetivos establecidos. La propuesta curricular y la actividad docente se deberían organizar, estructurar y caracterizar alrededor y dependiendo de este nuevo elemento.

4.2. Cambios en los contenidos objetos de evaluación

Posiblemente, el cambio más profundo se ha producido con referencia a la naturaleza de los aprendizajes. La calidad de un aprendizaje ya no se basa exclusivamente en el hecho de conocer más sobre un dominio concreto, sino en nuestra capacidad de utilizar de forma holística nuestros conocimientos, habilidades y actitudes con el fin de aplicarlos de manera activa y eficiente sobre tareas específicas. Con todo eso nos referimos al proceso que ya hemos descrito profusamente como desarrollo competencial y el problema radica en el enfoque de los procesos de evaluación sobre este nuevo tipo de aprendizaje.

4.3. Cambios en la lógica de la evaluación

Finalmente, el tercer gran cambio hace referencia a la nueva lógica que orienta los procesos evaluadores. La evaluación educativa, históricamente, se había centrado en el control de los resultados del aprendizaje. Posteriormente desplazó su preocupación a los procesos de petición de responsabilidades (accountability), lo cual significa implicar a toda la comunidad educativa en la responsabilización de la consecución de la calidad de los procesos y los resultados educativos. Es básicamente en esta última década cuando se descubre el enorme potencial de la evaluación como herramienta para gestionar los mismos aprendizajes y garantizar la calidad de los mismos. Se establece definitivamente la importancia de asociar los procesos evaluadores a los de desarrollo y potenciación de nuestra capacidad para aprender.

Es en este contexto en el que cobra carta de naturaleza la expresión evaluación continuada que subsume los tres conceptos contenidos en la ya superada tipificación de evaluación diagnóstica, sumativa y formativa. Bajo esta nueva acepción se entiende que la distinción entre evaluación y aprendizaje es puramente académica, en la práctica suponen dos tipos de actividades que comparten la misma intencionalidad. Toda actividad de aprendizaje debe ser evaluada y toda actividad evaluativa no es sino una actividad de aprendizaje de lujo.

La acción evaluativa confiere al aprendizaje de una tercera dimensión en el desarrollo del saber competente, la que te permite tomar conciencia no tan sólo de lo que sabes y de lo que no sabes sino de los elementos claves sobre los que se asienta los procesos de adquisición del conocimiento y te permiten desarrollar tu capacidad autónoma de aprender. Por el término "evaluación continua" no hemos de deducir que se trata de estar continuamente evaluando sino de entender que aprendizaje y evaluación forman parte del mismo "continuum"

Además, hay que señalar que la evaluación de las competencias alcanzadas por el estudiante ni tan sólo tiene el punto de vista de la evaluación de los resultados individuales del aprendizaje, sino que también adopta el punto de vista institucional, es decir, la calidad de una institución está asociada al grado en que consigue que sus estudiantes sean competentes en aquello descrito en el perfil de formación.

5. EVALUACIÓN Y APRENDIZAJE Y DESARROLLO COMPETENCIAL

Una vez analizado el concepto de competencias, su impacto en el currículo y apuntar algunos cambios en los enfoques evaluativos, entendemos que el aprendizaje y evaluación del desarrollo competencial precisa para su desarrollo la concreción de los siguientes elementos:

- Identificación de las competencias, fijación del nivel de desarrollo competencial y establecimiento de los niveles de logro.
- Descripción de las actividades donde se manifestará la competencia.
- Instrumentos o medios para evaluar el nivel de desarrollo.
- Estándares o criterios por los que se juzga si alguien es o no competente.

5.1. Identificación de las competencias y fijación del nivel de desarrollo competencial y establecimiento de los niveles de logro

Identificar y definir las competencias y fijar en su "continuum" el nivel de desarrollo deseable es importante con el fin de comunicar a los estudiantes qué se pretende alcanzar con el proceso de enseñanza-aprendizaje y en qué medida sus experiencias de aprendizaje y sus esfuerzos están dirigidos a esta consecución.

En la descripción de la competencia se tienen que señalar tanto los contenidos implicados como el nivel de complejidad del contexto en el que se tendrá que aplicar la competencia. Por eso, consideramos que la formulación de la competencia requiere un *verbo activo*, que identifique una acción que genere el resultado visualizable. De esta manera, hay que evitar el uso de los verbos como *conocer*, *entender* o *comprender* y utilizar otras formas verbales como *describe*, *identifica*, *reconoce*, *clasifica*, *compara*, *evalúa* o *valora*, *formula*, *argumenta*, *calcula*, *planifica*, *diseña*, etc. Además se requiere la descripción del objeto de la acción y el campo disciplinario en el que se fundamenta y el nivel de desarrollo que se desea alcanzar.

La fijación de los niveles de logro implica operativizar los de desarrollo de la competencia y con ello nos obligamos a establecer el nivel de contenido a aprender en relación a un contexto de realidad concreto. Con ello hemos de ser conscientes que estamos sustituyendo el término contenido por el de logro (enfaticando con ello la conversión de contenidos en aprendizajes). Este cambio conceptual pone en crisis los procedimientos tradicionales de evaluación incapaces de capturar este nuevo tipo de información evaluativa.

5.2. Descripción de las actividades donde se manifestará la competencia

La descripción de las actividades consiste en describir con precisión el tipo de actividad en la que se manifestará la competencia y los objetivos que se persiguen llevándola a cabo. Consiguientemente, se tienen que explicitar las competencias asociadas a esta actividad, es decir, qué conocimientos o habilidades llevan implícitos y en qué contextos se aplicarán, como también el nivel de profundidad o complejidad en el que se tendrían que concretar. Una vez definidas hay que describir qué competencias están implicadas en la actividad, en qué nivel y contexto se trabajarán y de qué medios se dispondrá, se pueden concretar los resultados de aprendizaje esperados en cada actividad, es decir, sus resultados observables. De este modo será posible establecer qué tipo de evidencias se producen y cómo se pueden recoger con el fin de analizar el nivel de consecución de las competencias descritas. Este nivel de descripción es necesario en las actividades que son objeto de evaluación; por lo tanto no hay que hacerlo de una manera tan detallada para el resto de actividades de aprendizaje, donde se pueden introducir competencias que no sean objeto de evaluación.

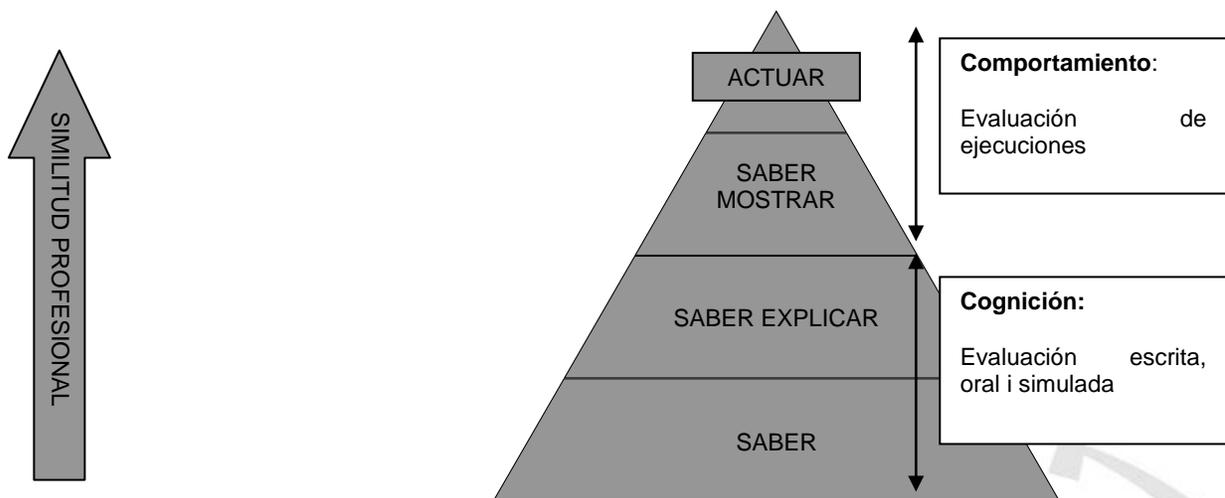
5.3. Instrumentos o medios para evaluar la competencia

La determinación del tipo de instrumento que hay que aplicar para la recogida de evidencias depende fundamentalmente de la naturaleza del resultado de aprendizaje que se tiene que capturar. Si bien, tal como hemos visto en la figura 1, la competencia sólo se puede evaluar en la acción, para poder adquirirla hace falta haber alcanzado previamente una serie de conocimientos, habilidades y actitudes que habremos descrito bien de acuerdo con los resultados de aprendizaje o en términos de objetivos, según si nuestra perspectiva es lo que pretende el profesor o bien lo que tendrá que demostrar el estudiante. La pirámide de Miller (1990) puede ser una manera útil de ayudar a escoger estrategias de evaluación

coherentes con resultados de aprendizaje descritos por el profesor. Así, se puede evaluar sólo el hecho de saber (por ejemplo, por medio de una prueba tipo test) o el hecho de saber explicar, que ya requiere una gestión del conocimiento adquirido; o bien se puede plantear una simulación en la que el estudiante actúe en situaciones controladas; y, finalmente, hay que demostrar en actuaciones la adquisición de una competencia.

En general se entiende que la acción evaluativa supone desde la perspectiva moderna el uso integrado de diferentes técnicas de recogida de información (Stake, 2008), aunque evidentemente la comprobación final de la asunción de un nivel concreto de desarrollo de una competencia exige necesariamente el uso de pruebas basadas en ejecuciones.

FIGURA 2. LA PIRÁMIDE DE MILLER (1990)



La pirámide distingue dos grandes tipos de pruebas que podríamos clasificar en *evaluación tradicional* (o pruebas de papel y lápiz) y *evaluación de ejecuciones*.

La evaluación tradicional engloba lo que podríamos denominar las típicas pruebas, en las que se hace más hincapié en los objetivos de conocimientos y de saber. Dentro de la evaluación tradicional hay que distinguir entre unas pruebas que enfatizan habilidades de bajo orden (recuerdo, comprensión), mientras que hay otras que enfatizan el pensamiento de alto orden (aplicación, síntesis, evaluación).

Mediante la aplicación de las estrategias y los instrumentos del enfoque tradicional, no sabemos a ciencia cierta cómo están aprendiendo nuestros estudiantes y como les afecta la aplicación del currículum, por todo ello surgen movimientos que enfatizan el uso de métodos que faciliten la observación directa del trabajo de los estudiantes y de las habilidades y actitudes que ponen en juego. Este nuevo enfoque se conoce como el de la evaluación basada en ejecuciones.

La evaluación de ejecuciones es muy variada y permite abarcar un rango mucho más amplio de competencias, sea de habilidades disciplinarias o de competencias transversales. Exige en definitiva que

los estudiantes, en un momento determinado, demuestren, construyan, o desarrollen un producto o solución a partir de unas condiciones definidas previamente (Rodríguez, 2000).

Khatti y Sweet (1996) señalan que adoptar la evaluación basada en ejecuciones implica los siguientes pasos a efectuar por los estudiantes:

- Estructurar las tareas objeto de evaluación.
- Aplicar una información previa.
- Construir respuestas.
- Explicar el proceso que les ha llevado a una determinada respuesta.

Es evidente que las tareas reales incluyen todas las condiciones contenidas en el listado anterior, la realidad aporta autenticidad a las pruebas, pero es del todo ingenuo pensar que seremos capaces de generar tantos escenarios de realidad como necesidades evaluativas tengamos.

Por todo ello cabe señalar a los educadores que tiendan a la ideación de aproximaciones con el mayor carácter de autenticidad posible. Frazee y Rudnitske (1995) nos señalan como aproximaciones a la autenticidad: La resolución de problemas, el diseño y la ejecución de experimentos sobre problemas reales, la creación de videos, trabajos de campo, desarrollo de demostraciones, la construcción de modelos, la promoción de debates, la publicación de diarios escolares, simulaciones por ordenador, etc., por supuesto que nos se trata de una lista exhaustiva.

Sin embargo la clave está en pensar que dado que no es posible (seguramente, ni deseable) mantener de forma absoluta la autenticidad de todas las actividades escolares que propongamos, lo que sí es importante es que la evaluación se produzca bajo situaciones estratégicas, que conecten a los estudiantes con las situaciones y las condiciones de la vida real, o expresado en otros términos, que el reto cognitivo que suponga la realización de la actividad sea del mismo nivel y dificultad que el que se encontrarían en una situación real.

Podríamos resumir con Wolf y Reardon (1996), las características más básicas del enfoque basado en ejecuciones en las siguientes:

- Muestras: Experimentos de los alumnos, debates, portafolios, productos de los estudiantes.
- Juicio evaluativo: Basado en la observación, en la subjetividad y en el juicio profesional.
- Focalización; De manera individualizada sobre el estudiante a la luz de sus propios aprendizajes.
- Objetivo 1: Habilita al evaluador a crear una historia evaluativa respecto del estudiante o del grupo.
- Objetivo 2: Provee información evaluativa de manera que facilita la acción curricular.
- Permite a los estudiantes participar en su propia evaluación.

5.4. Estándares o criterios por los que se juzga si alguien es o no competente

Finalmente, tenemos que establecer los criterios valorativos que nos permiten emitir los juicios de valor respecto de los resultados alcanzados. Si aplicamos los criterios de evaluación sobre los resultados de aprendizaje, podemos expresar estos resultados en términos de estándares de ejecución. Aquí no tan sólo

expresamos lo que tiene que hacer, sino que también establecemos los niveles de ejecución que permiten establecer juicios con respecto al nivel de logro del aprendizaje.

Además, si queremos mejorar, de manera significativa, la precisión de nuestros juicios valorativos y, consiguientemente, la constancia de las valoraciones emitidas con respecto a una misma ejecución (especialmente cuando se hacen por parte de diversos evaluadores), antes hay que aclarar los aspectos o las dimensiones que se quieren evaluar, como también los indicadores o las evidencias que identifican los niveles de valoración que proponemos. Para conseguir esta aclaración es conveniente utilizar ejemplos de lo que pretendemos y, para su buen funcionamiento, tendrían que estar insertados en el marco de un esquema general de evaluación.

Y para acabar, se tiene que proceder al análisis de toda la información de evaluación con respecto a cada uno de los resultados evaluadores en el nivel de exigencia esperado y determinar si se han alcanzado y con qué grado de suficiencia todas y cada una de las competencias que llevaba implícita la realización de la actividad.

6. A MODO DE SÍNTESIS

Tal como hemos visto los procesos de aprendizaje orientados al desarrollo competencial implican un cambio cultural en profundidad por parte de todos los agentes que operan en el sistema educativo. Incorporar las nuevas formas de relacionarnos con el conocimiento y trasladar al sistema educativo sus consecuencias constituye el gran reto del siglo XXI.

Todo ello nos sitúa en una nueva lógica del sistema que para gestionarlo y hacerlo viable es preciso ampliar el campo de acción de la educación, incorporando nuevos agentes, estableciendo como necesidad la relación con el entorno, promoviendo el trabajo en red de los centros, etc. Aunque posiblemente los dos extremos más prioritarios son modificar las propuestas curriculares integrando las competencias y dinamizando su formato y estableciendo y regulando la necesaria y estrecha vinculación entre aprendizaje y evaluación.

En relación a este último la Comisión on Higher Education señala como pasos a seguir los siguientes (CHE, 1997).

- Identificar los conocimientos y las habilidades que los estudiantes deberían aprender y las competencias que deberían desarrollar para estar capacitados para su inserción activa en la vida y en la sociedad como personas y ciudadanos.
- Identificar los atributos personales que los estudiantes deberían adquirir y/o desarrollar en una institución educativa.
- Diseñar propuestas curriculares de carácter integral y dinámico que estructure formal y lógicamente todos estos elementos y establezca los nexos funcionales básicos para favorecer su operativización y plasmación en los programas formativos
- Considerar y decidir respecto del conjunto de procedimientos e instrumentos evaluativos que habrá que aplicar de forma estratégicamente combinada para realmente capturar información relevante respecto del rendimiento académico y del desarrollo personal.

- Recoger y analizar la información evaluativa requerida para juzgar la efectividad de la institución en lo que hace referencia a la enseñanza y al aprendizaje.
- Desarrollar un sistema eficiente para comunicar la información de forma que facilite la mejora de la enseñanza y del aprendizaje.

Pero no se trata simplemente de una nueva reestructuración de los posicionamientos clásicos, sino que lo que perseguimos es un cambio en la naturaleza, en la lógica de actuación y en el "rol" de los agentes y elementos que intervienen.

Así cuando hablamos de conocimientos, habilidades y competencias, es fundamental entender que hacemos referencia a aquellos que son también socialmente relevantes y no basta con que lo sean desde el punto de vista académico.

Cuando mencionamos el desarrollo de los atributos personales, queremos señalar la necesidad de dinamizarlos intencionadamente y no dejar su activación al azar. Hay que preverlo desde el inicio de la acción educativa y proceder también a su evaluación.

La incorporación del desarrollo competencial como objetivo básico del aprendizaje nos obliga a ampliar la tipología de los procedimientos evaluativos empleados y a pensar en ellos como un todo estratégicamente combinado e integrado, deberemos necesariamente incorporar, al menos para los momentos más nucleares en la adquisición de los aprendizajes, el uso de las pruebas basadas en ejecuciones.

Será preciso incorporar nuevas formas organizativas en las instituciones educativas y desarrollar el sentido colaborativo entre los educadores.

Convendrá modificar los sistemas de gestión de los centros y para ello se precisarán estructuras y sistemas de funcionamiento más ágiles y eficientes tanto para la recogida de la información, como para su almacenamiento y especialmente para su explotación y uso. Todo ello debería conllevar precisión y objetividad en la comunicación de la información, rapidez en sus aplicaciones y flexibilidad y eficacia en la activación de los procesos de cambio.

Fundamentalmente el colectivo más afectado por los cambios que afectan a los procesos de aprendizaje y de evaluación es sin duda el del profesorado. Esta realidad obligará a las administraciones y a los centros a incrementar los procesos de sensibilización y formación respecto al manejo de las competencias en el contexto del diseño curricular y a como trasladar su desarrollo al programa.

Todo ello nos va a obligar en palabras de Haergreaves (1999) a reconstruir nuestro pensamiento en el que hace referencia al significado del aprendizaje y la evaluación y a incorporarnos a la cultura fijada por el nuevo paradigma.

REFERENCIAS BIBLIOGRÁFICAS

Barbarà, E., Rodera, A.M., Aguado, G., Guàrdia, L. y Vall-llovera, M. (2009). Evaluation of Learning Base don Transverse Competences using Web 2.0 Orientations: The e-Transfolio. *8th European Conference on e-learning*. Bari: Italy.

- Barrón, C. (2000). La evaluación basada en competencias en el marco de los procesos de globalización. En Valle, M. (coord.) (2000). *Formación de competencias y certificación profesional*. México: Centro de Estudios de la Universidad, UNAM.
- Borghesi, M. (2005). *El sujeto ausente. Educación y escuela entre el nihilismo y la memoria*. Madrid: Fondo Editorial.
- Cano, E. (2008). La evaluación por competencias en la educación superior. *Profesorado: Revista de curriculum y formación del profesorado*, 12(3), pp.1-16.
- Eurydice (2003). *Key competences: A developing concept in general compulsory education*. Accesado el 20 de junio de 2010 en <http://eurydice.org>.
- Frazeo, B. y Rudnitski, R.A. (1995). *Integrated Teaching Methods*. Albany (NY): Delmar Publishers.
- González, O. (2002) Evaluación Basada en Competencias. *Revista de Investigación*, 53, pp. 11-31.
- Hargreaves, A. (1999). *Four Perspectives on Classroom Assessment Reform*. B.
- Hutmacher, R.W. (1999). L'avaluació en la transformació de les modalitats de govern als sistemes educatius. *Tendències europees en avaluació i educació*. Barcelona, Generalitat de Catalunya: Consell Escolar de Catalunya.
- Khattari, N. y Sweet, D. (1996). Assessment Reform: Promises and Challenges. En M.B. Kane y R. Mitchell (Eds.). *Implementing Performance Assessment*, pp. 1-21. Mahwah (NJ): Lawrence Erlbaum.
- Lasnier, F. (2000). *Reussir la formation par compétences*. Mont-real: Guérin.
- Le Boterf, G. (2001). *Ingeniería de las competencias*. Barcelona: Gestión.
- Mateo, J., Escofet, A., Martínez, F. y Ventura, J. (2009). *Guía para la evaluación de competencias en el trabajo de fin de grado en el ámbito de las ciencias sociales y jurídicas*. Barcelona: Generalitat de Catalunya. Agència per a la Qualitat del Sistema Universitari a Catalunya.
- Mateo, J. y Martínez, F. (2008). *Medición y Evaluación Educativa*. Madrid: La Muralla.
- Miller, G. (1990). The assessment of skills/competencies/performance. *Academic Medicine (Supplement)*, 65, pp. 63-67.
- Morin, E. (2001). *Los siete saberes necesarios para la educación del futuro*. Barcelona: Seix Barral.
- NCES (2002). *Defining and Assessing Learning: Exploring Competency-Based Initiatives*. National Postsecondary Education Cooperative. Accesado el 2 de Septiembre de 2010 en <http://nces.ed.gov/publicsearch>.
- Perrenoud, P. (2004). *Diez nuevas competencias para enseñar*. Barcelona: Graó.
- Rodríguez, M^a.L., Serreri, P. y Del Cimmuto, A. (2010). *Desarrollo de competencias: Teoría y Práctica*.
- Rodríguez, S. (2000). La evaluación del aprendizaje de los alumnos. Ponencia presentada al *I Congreso Internacional: Docencia Universitaria e Innovación*. Barcelona: Institut de Ciències de l'Educació.
- Roe, R.A. (2002). What makes a competent psychologist? *European Psychologist*, 7(3), pp. 192-202.
- Rychen, D. y Salganik, L. (2005). *Defining and Selecting key competencies. Executive Summary*. Accesado el 15 de Julio de 2010 en <http://oecd.org/dataoecd/47/61/35070367/pdf>.
- Stake, R.E. (2008). La ventaja de los criterios, la esencialidad del Juicio en *Revista Iberoamericana de Evaluación educativa*. 1(3e), pp.18-28.

- Trilling, B. y Fadel, Ch.(2009). *21st. century skills. Learning for life in our times*. S.Francisco,CA: Jossey-Bass.
- Weinert, F.E. (2001). Concept of Competence. A conceptual clarification. En D. Rychen y L. Salganik (Eds) *Defining and selecting keycompetencies*, pp. 45-66. Kirkland: Huber-Hogrefe Publishers.
- Wolf, D.P. y Reardon, S.F. (1996). Acces to Excellence through New Forms of Student Assessment. En J.B. Baron y D.P. Wolf (Eds). *Performance Based Student Assessment: Challenges and Possibilities*. Ninety-fifth Yearbook of the National society of education, Part I. Chicago (IL.): University of Chicago Press.



O EFEITO DA CAPACITAÇÃO DOCENTE NO DESEMPENHO DOS ALUNOS: UMA CONTRIBUIÇÃO PARA A FORMULAÇÃO DE NOVAS POLÍTICAS PÚBLICAS DE MELHORIA DA QUALIDADE DA EDUCAÇÃO BÁSICA

THE EFFECT OF TEACHER TRAINING IN THE PERFORMANCE OF STUDENTS: A CONTRIBUTION TO THE FORMULATION OF NEW POLICIES TO IMPROVE THE QUALITY OF BASIC EDUCATION

Nilma Santos Fontanive y Rubén Klein

Revista Iberoamericana de Evaluación Educativa 2010 - Volumen 3, Número 3

<http://www.rinace.net/riee/numeros/vol3-num3/art04.pdf>

Fecha de recepción: 18 octubre de 2010
Fecha de dictaminación: 23 octubre de 2010
Fecha de aceptación: 25 octubre 2010

Na última década a preocupação com a qualidade dos professores está presente na agenda mundial, principalmente pelas evidências dos baixos desempenhos demonstrados pelos alunos de quase todos os sistemas escolares no mundo. Por exemplo, o Programa Internacional de Avaliação de Alunos-PISA, que congrega os países membros da OCDE e outros por adesão, na sua edição de 2006 avaliou o desempenho de alunos de 15 anos de 57 países e seus resultados atestam o fracasso de grande parte das nações em prover uma educação de qualidade para sua população de estudantes.

Entretanto é importante acrescentar que, ainda hoje, não há consenso entre educadores sobre o fato da qualidade docente ser medida pelos resultados da aprendizagem dos alunos e mais, que o desempenho dos alunos avaliados por testes padronizados seja uma medida da qualidade docente, embora numerosos estudos venham buscando estabelecer estas relações

Segundo Laura Goe (2007) as pesquisas que fazem a ligação entre qualidade docente e resultados de desempenho dos alunos podem ser sintetizadas em três categorias. Tais pesquisas buscam respostas para as seguintes questões: o que é qualidade docente e como ela pode ser medida? Quão importante é a qualidade docente para a aprendizagem dos estudantes? Qual o valor da experiência docente para a aprendizagem dos alunos?

Este artigo vai apresentar os resultados da pesquisa realizada com professores da Rede Estadual do Tocantins que participaram de atividades de capacitação, com vistas a melhorar a qualidade docente, conduzidas pelos autores do artigo durante 6 anos, e que tiveram impacto no desempenho dos alunos medido pela Avaliação Nacional da Educação Básica (SAEB) - Prova Brasil¹.

1. A QUALIDADE DOCENTE: CONCEITUAÇÃO E QUADRO DE REFERÊNCIA PARA ANÁLISE

Analisando algumas pesquisas que investigam a qualidade docente vê-se que elas podem enfocar seus inputs, os processos e os resultados de aprendizagem dos alunos.

Neste artigo será priorizada a discussão sobre a qualidade docente como efetividade docente, definida como a contribuição demonstrável para o crescimento da aprendizagem do aluno por ser este o objetivo central do estudo realizado com professores e alunos do Ensino Fundamental da rede estadual do Tocantins no período de 2003 a 2008;

O estudo buscou investigar em que condições o Programa de Capacitação de Professores pode causar efeitos positivos no desempenho dos alunos, medidos por avaliações nacionais externas, procurando descobrir quais os componentes desse Programa produziram o aumento das médias dos alunos nas disciplinas de Matemática e Língua Portuguesa.

Antes de prosseguir, deve-se apontar o caráter não consensual do conceito de efetividade docente considerado apenas na sua relação com ganhos de aprendizagem cognitiva, pois bons professores podem acrescentar também outros valores, tais como, motivação e engajamento dos estudantes para adquirir novos conhecimentos, habilidades e espírito de colaboração com os colegas. Portanto, a aprendizagem

¹ O SAEB/PROVA BRASIL é uma avaliação nacional do universo dos alunos de 4ª e 8ª séries do Ensino Fundamental, das escolas públicas urbanas brasileiras, com pelo menos 20 alunos na série, nas disciplinas de Língua Portuguesa e Matemática, realizada a cada 2 anos. As Escalas do SAEB/PROVA BRASIL vêm permitindo acompanhar o desempenho dos alunos desde 1995.

dos alunos pode ser avaliada de diferentes maneiras e não apenas através dos seus resultados nos testes e exames de final de curso. Independentemente dos instrumentos usados como medida, a maioria dos educadores e autoridades educacionais concorda que há efetividade docente quando a aprendizagem dos estudantes aumenta (Darling-Hammond, 2007).

Focalizando a efetividade docente em termos dos resultados dos estudantes, importa saber quais são os professores que estão obtendo resultados e porque estão obtendo. Quando os dados de desempenho dos alunos estão disponíveis, esses podem ser usados como *proxy* da efetividade docente, embora seja recente a aceitação nos meios acadêmicos de que o desempenho dos alunos seja um componente da medida da qualidade docente. Até há bem pouco tempo atrás, educadores e gestores avaliavam a qualidade docente usando somente critérios de qualificação, de certificação, anos de experiência profissional, aquisição de graus mais avançados de educação, como mestrado e doutorado, por exemplo.

Essa qualificação não deve ser desprezada nos estudos do efeito - professor, pois ela serve como um controle de qualidade do desempenho docente e pode algumas vezes ser uma boa preditora do sucesso dos alunos. Entretanto, nos dias atuais há um consenso quanto à qualificação do professor ser uma condição necessária, mas não suficiente para produzir efeitos positivos no desempenho dos alunos.

A análise da literatura revela também que a ciência para medir a qualidade docente ainda está em estágios de desenvolvimento. Não se pode ainda contar com respostas, ou achados definitivos para todas as questões nas inúmeras pesquisas da área que, embora em desenvolvimento, já há um bom número delas que permite manter acesa a discussão.

A qualidade docente compreende vários aspectos que faz um professor ser “bom”. Entre eles, Goe (2007) inclui: a qualificação, formação, capacidade, expertise, caráter, performance e sucesso profissional. Ainda segundo a autora, há numerosos instrumentos para avaliar mais do que uma dimensão da qualidade docente e o uso de instrumentos variados requer um significativo investimento de tempo para desenvolvê-los e validá-los, treinar pessoas para aplicá-los, constituir os bancos de dados e analisar seus resultados

Voltando a Laura Goe (2007, p.1), a qualidade docente pode ser definida de diferentes maneiras e muito embora não haja consenso sobre uma definição única, há uma relativa concordância entre os pesquisadores da área que a qualidade docente pode ser evidenciada em professores com as seguintes características:

- Qualificações e experiência apropriada para a matéria e o nível de ensino onde atua;
- Altas expectativas para os estudantes, especialmente para os de baixo desempenho;
- Criação de um ambiente na sala de aula que encoraje a participação de todos os estudantes nas atividades de aprendizagem propostas;
- Desejo de ajudar os estudantes a atingirem altos níveis de proficiências;
- Habilidade em monitorar novos professores e ação estabilizadora da permanência e coesão do corpo docente na escola;
- Disposição para trabalhar com afinco para ajudar os estudantes com necessidade de maior apoio, mesmo que os resultados do desempenho desses alunos não reflitam a contribuição do professor.

Algumas dessas características que expressam a qualidade docente foram enfatizadas no Programa de Capacitação de Professores desenvolvido pelos autores do artigo.

As pesquisas sobre os desempenhos dos alunos em exames padronizados, associadas às características das escolas onde eles estudaram e as dos seus professores, não são recentes. Hanushek (1971) foi um dos primeiros pesquisadores a isolar de maneira analítica as diferenças de desempenho entre classes de uma escola e mostrou que existem diferenças de progresso bastante significativas segundo as salas de aula que os alunos frequentam. Na obra citada, o autor, e, posteriormente Veldman e Brophy (1974), concluem, por inferência, que os efeitos de variabilidade do desempenho entre salas de aula eram produzidas pelos professores. Hanushek, entretanto, assinala que o efeito sala de aula não parece ser dissociado de um possível efeito de composição dos alunos daquela classe.

Baseados nessas premissas foram criados numerosos programas que tentam colocar em ação práticas e processos de ensino que se mostraram ligados às boas performances e, em seguida, novas correntes de pesquisa direcionam-se para avaliar se a implementação de tais programas conduziram a melhoria do desempenho dos alunos e das escolas, ou seja, origina-se um conjunto de pesquisas ligadas ao que convencionou-se chamar de " *School Improvement*". (Reynolds *et al.*, 1993).

2. A EFETIVIDADE DOCENTE E A SUA MEDIDA

A efetividade docente se tornou um tema de grande interesse entre os pesquisadores que estudam qualidade docente. Conforme assinalado anteriormente, essa área de investigação ainda necessita de mais trabalhos para demonstrar quais são as características docentes, qualificações, práticas que contribuem para as diferenças de desempenho encontradas nos alunos.

Outro aspecto importante na investigação na área da efetividade docente é a questão metodológica, pois o desenho das pesquisas deve possibilitar fazer a ligação entre o professor (práticas, características docentes e formação) com o desempenho dos alunos, turmas e escolas.

Aaronson, Barrow and Sanders (2003), por exemplo, conduziram um estudo unindo os dados do Chicago Public Education High School de professores e alunos de 8º e 9º anos em Matemática. Usando o método de valor agregado os autores encontraram que professores de alta qualidade acadêmica (médias de desempenho com 2 desvios padrão acima da média geral) adicionam entre 25% a 45% no crescimento da média de um ano de escolaridade nos escores do estudante em Matemática. Os autores tentaram correlacionar alguns dados disponíveis dos professores, idade, experiência, certificação e diploma de graduação na área e encontraram pouca correlação entre os fatores observáveis de professor e o desempenho dos alunos, exceto a graduação em Matemática. Eles concluíram que, o que os professores de alta qualidade fazem nas salas de aula é mais importante do que suas qualificações iniciais. Para os autores, o método de valor agregado não explica porque os professores variam de qualidade e nada se sabe do que se passa na sala de aula que ajude a prever com quais professores os alunos ganharão mais. Há ainda algumas questões metodológicas que precisam também ser contornadas, pois, de acordo com Braun (2004) é difícil isolar o efeito do professor de outros efeitos do nível sala de aula tais como clima, influência dos pares, disponibilidade de livros e outros materiais didáticos e ainda de outros fatores escolares que podem contribuir para a aprendizagem dos estudantes e que estão fora do controle do professor.

Noell (2006) usou o método do valor agregado nos escores obtidos pelos alunos para examinar a eficácia do professor. Na primeira fase da pesquisa os escores do valor agregado foram calculados para estudantes de 4º a 9º anos em 66 das 68 escolas públicas de Louisiana e ligados aos professores. Sem surpresa, encontraram que o maior preditor de ganhos de aprendizagem dos alunos foi o escore obtido pelo professor no teste inicial na área que ensinavam, Matemática ou Linguagem. Na segunda fase o autor relacionou professores de um programa de capacitação e estes foram identificados e ranqueados de acordo com estimativas de efetividade. Foram encontradas relações positivas entre os dois.

Um dos mais importantes estudos ligando os impactos das escolas e dos professores ao desempenho dos estudantes foi realizado por Rivkin, Hanushek, and Kain, em 2005. Usando dados do Texas, os autores examinaram os resultados obtidos pelos alunos no Texas Assesment of Academic Skills (TAAS) e componentes observáveis (educação do professor e experiência) e componentes não observáveis (residuais). Focalizando alunos de 3º ao 7º anos e escores de mais de 140 000 a 455 000 alunos (os números variavam dependendo do ano e série), os autores encontraram que as características observáveis têm um pequeno mas significativo efeito nos ganhos dos alunos, embora a maior efetividade docente está ligada a diferenças não observáveis da qualidade da instrução. As principais conclusões dos autores contrariam as divulgadas pelo Relatório Coleman, pois afirmam que as escolas e os professores importam para os resultados dos alunos. O modelo da pesquisa e os dados usados permitiram aos autores chegar a 4 importantes conclusões:

- Não há evidências quanto à obtenção de um diploma de mestrado aumentar as habilidades docentes;
- Os ganhos na qualidade docente são grandes no primeiro ano de exercício e menores nos anos seguintes, em até 3anos. Há pequenas evidências que os ganhos na qualidade docente continuem após os três primeiros anos;
- O tamanho das classes têm um modesto, mas estatisticamente significativo efeito em matemática e leitura nas séries iniciais, mas esses efeitos decrescem conforme aumenta a progressão nas séries escolares;

Entretanto o aspecto mais polêmico dos resultados da pesquisa é a afirmação que, embora reconhecendo que o professor tenha um poderoso efeito no desempenho dos alunos em Leitura e em Matemática, somente uma parte pequena da variação da qualidade docente é explicada por características observáveis como a sua educação ou experiência. Esta conclusão tem um grande impacto nas políticas públicas de formação e educação continuada de professores, uma vez que são as características que não têm sido observadas que explicam a maior variação das proficiências dos alunos.

Espera-se que os resultados obtidos no estudo com os professores possibilitem acrescentar novos elementos às conclusões da pesquisa conduzida por Hanushek, pois foram encontrados efeitos positivos, de certa magnitude, relacionando a capacitação docente e melhoria de desempenho dos alunos

No Projeto de Capacitação dos Professores do Tocantins, cujos efeitos estão sendo discutidos nesse artigo houve uma grande ênfase em capacitar o professor no conteúdo a ser ensinado, na metodologia do ensino e nas práticas na sala de aula.

Os estudos têm consistentemente encontrado que professores efetivos provocam impactos mais significativos na aprendizagem dos estudantes do que qualquer outro fator, e que professores efetivos

podem compensar desafios tais como pobreza, etnia e lacunas de aprendizagem (Rivikin *et al.*, 2002; Clotfelter *et al.*, 2007).

A magnitude da diferença nos desempenhos dos alunos como efeito do ensino, faz com que muitos pesquisadores venham enfatizando que a efetividade do professor, mais do que qualquer outro indicador da qualidade docente, deve ser a área em que os educadores e os implementadores de políticas públicas de melhoria da qualidade da educação precisam focar sua atenção se o objetivo é o de aumentar os ganhos de aprendizagem dos alunos (Alliance for Excellent Education, 2008, p. 2).

Apesar das dificuldades metodológicas apontadas, alguns efeitos foram consistentemente encontrados nos diferentes estudos sobre a qualidade docente. O primeiro deles é a correlação positiva entre a Graduação do professor e certificação ou aprovação em exames de Matemática e o desempenho dos alunos nessa disciplina, particularmente no ensino secundário e médio. Em outras áreas curriculares, como Linguagem, Estudos Sociais e Ciências, não há tantos estudos focalizando os aspectos que as pesquisas em Matemática vêm estudando. Alguns pesquisadores sugerem que a aprendizagem da Matemática se dá em grande medida na escola, ao contrário de outras disciplinas, como a língua, por exemplo, que são influenciadas por outras situações sociais não mediadas pela escola. Nesse raciocínio, os efeitos do professor são mais sensíveis de serem encontrados nessa área curricular e para desenvolver habilidades de alta ordem é fundamental que os professores de Matemática tenham competência para guiar seus alunos nas atividades práticas na sala de aula que conduzem a descobertas.

Os estudos sugerem ainda que seria recomendável aumentar as exigências de ingresso no magistério para ensinar Matemática, inclusive oferecendo pagamentos diferenciados para atrair e manter professores bem treinados e certificados na carreira docente.

O segundo efeito encontrado pelas pesquisas refere-se ao número de anos de exercício profissional, evidenciando que o pico de performance se dá até o quinto ano de magistério. Assim torna-se bastante relevante traçar políticas efetivas para distribuir equitativamente os professores pelas escolas, evitando-se alocar professores inexperientes em escolas com altas porcentagens de alunos de nível socioeconômico mais desfavorecidos ou de minorias raciais.

Outro aspecto importante a considerar diz respeito ao contexto do ensino. Até que ponto o ambiente social da escola interfere no desempenho de alunos e professores? Essa é uma questão crucial para promover-se a equidade. A pergunta que precisa ser respondida é: dado um determinado contexto que características devem ter os professores para produzir um efetivo impacto nos ganhos de aprendizagem dos alunos? Ou que práticas os professores efetivos das escolas que atendem a alunos em situação de vulnerabilidade social adotam para levar esses alunos a altos níveis de desempenho?

O que importa é que as definições da qualidade docente dependem dos propósitos que se tem em mente e essas definições devem, em princípio, englobar dois componentes:

- 1) um conjunto de qualificações ligadas fortemente à disciplina e ao nível de ensino tem que estar garantidas antes que o professor assuma uma classe e;
- 2) algum mecanismo para avaliar a efetividade docente capaz de prever altos níveis de aprendizagem de seus alunos.

A combinação das duas definições conduz a dois estágios de avaliação da qualidade docente: o primeiro, baseado nos diplomas e certificações dos professores e alguma medida de efetividade a ser obtida antes

que o professor assuma seu posto com uma turma. Essa avaliação envolve a combinação de provas ou certificações, de avaliação de especialistas ou de pares, portfólios docentes. O segundo é avaliar o professor já em exercício pelos escores do desempenho dos seus alunos, medidos por avaliações externas.

3. O PROGRAMA DE CAPACITAÇÃO DE PROFESSORES DO ENSINO FUNDAMENTAL DO TOCANTINS

As atividades de capacitação dos professores de Língua Portuguesa e de Matemática do Ensino Fundamental iniciaram-se em junho de 2003 e ao longo deste período de 6 anos, o Programa de Capacitação sofreu diversas modificações no formato originalmente concebido, embora o principal eixo integrador da proposta permaneceu constante, associando a atualização dos conteúdos das disciplinas às práticas de sala de aula, mediante o desenvolvimento de metodologias e de material didático para professores e alunos que favorecessem mudanças na dinâmica das interações entre eles e os conteúdos de ensino.

As três orientações metodológicas que nortearam o programa de capacitação de professores naquele ano de 2003 permaneceram até o ano de 2008. Cada uma delas será discutida à luz de uma literatura selecionada sobre mudanças na prática docente que parecem produzir efeitos positivos na melhoria do desempenho acadêmico dos alunos.

A adoção do ensino participativo e o estímulo ao trabalho em grupo são recomendações recorrentes na literatura pedagógica desde o advento da Escola Nova surgido nos Estados Unidos, no final do século XIX, que postulava uma profunda reestruturação do ensino baseada em métodos ativos e experimentais.

Para aumentar a participação dos alunos no seu processo de aprendizagem o Programa de Capacitação propôs alterar a dinâmica da sala de aula, mudando o foco da atividade do professor para os alunos, organizados em grupos de trabalho voltados para a discussão e resolução de problemas ou exercícios desafiadores.

Uma das recomendações mais frequentes feitas por autores que pesquisaram os fatores que favorecem o ensino efetivo, nas últimas décadas, tem sido o de substituir o ensino frontal, unidirecional – no qual o ato de ensinar é quase exclusivamente desempenhado pelo professor enquanto os alunos passivamente ouvem ou copiam do quadro de giz – por situações que envolvam ativamente os alunos nas atividades de aprendizagem e que criem condições de troca de experiências entre eles (Bressoux, 2003).

Os processos de trabalho em grupo, em particular em grupos cooperativos e seus efeitos nas aquisições dos alunos, já foram bem estudados nos últimos 30 anos e uma grande quantidade de pesquisas sobre o tema pode ser encontrada na literatura.

Slavin, um dos teóricos mais importantes da aprendizagem cooperativa, no seu artigo *Cooperative Learning and Achievement: Research and Theory* (2003) faz uma síntese das pesquisas voltadas para estudar os efeitos do uso em sala de aula de grupos cooperativos nos desempenhos dos alunos. Segundo o autor, em mais de meio quarto de século, foram realizados centenas de estudos que compararam os efeitos da aprendizagem cooperativa utilizando variados métodos de controle, apoiados em inúmeras medidas de resultados obtidas nas diversas áreas curriculares e em todas as séries escolares. Os estudos abrangeram também diferentes cenários educacionais em muitos países.

Esses estudos produziram um conhecimento acumulado sobre os efeitos de vários tipos de intervenções cooperativas e dos mecanismos responsáveis por esses efeitos. Slavin ressalta, porém, que a aprendizagem cooperativa não é apenas um objeto de teoria e pesquisa, ela é usada, em alguma medida, por milhões de professores americanos, por exemplo, em um levantamento nacional realizado em 1993 (Slavin, op.cit., p.2) encontrou-se que 79% dos professores do Ensino Fundamental e 62% dos do Ensino Médio usam alguma técnica de aprendizagem cooperativa nas suas turmas. Os princípios da aprendizagem cooperativa também não se reduzem a um único método de ensino ou assumem um mesmo formato nas diferentes aplicações práticas no cotidiano das escolas, embora todos tenham um elemento comum: postularem uma mudança na atmosfera da sala de aula para aumentar a confiança dos alunos encorajando-os a participar, falar, debater, discordar e argumentar ao invés de esperar que o professor traga as soluções prontas ou explique os tópicos das disciplinas, sentados nas suas cadeiras, levantando a mão para perguntar e esperando serem chamados.

Para favorecer essa mudança da prática na sala de aula o Programa de Capacitação, conduzido pelos autores deste artigo, adotou duas estratégias principais: a elaboração de material didático desafiador para professores e alunos e a participação dos professores em oficinas com trabalho em grupo.

O material didático era composto de apostilas para o professor contendo orientações quanto o desenvolvimento de habilidades constantes dos guias curriculares editados pela Secretaria de Educação do Tocantins, sugestões de atividades a serem desenvolvidas com os alunos e, quando era o caso, alguns aportes teóricos sobre determinado tópico.

Para os alunos, os materiais assumiram o formato de roteiros de atividades ou exercícios para serem trabalhados e resolvidos em grupos.

Em nenhum momento pretendeu-se substituir os livros didáticos distribuído gratuitamente pelo Governo aos alunos brasileiros por esses materiais. Eles eram complementares aos livros que os professores adotavam nas suas escolas. Entretanto, uma pesquisa preliminar feita pelos autores deste artigo junto aos professores revelou que havia entre eles grande variedade de livros adotados e, assim, era preciso criar práticas unificadoras de abordagens e de conteúdos de ensino.

Acrescente-se também que ao longo desses 6 anos do Programa de Capacitação, os materiais didáticos sofreram numerosas alterações no seu formato, embora tenham conservado algumas características editoriais da primeira versão de 2003

A terceira e última orientação metodológica implementada pelo Programa de Capacitação de Professores foi o uso de avaliações formativas para acompanhar o progresso dos alunos e fornecer aos professores informações sobre as necessidades de re-ensino ou de reforço com suas turmas. Há um grande consenso entre os teóricos da aprendizagem sobre o fato de a avaliação periódica e constante ser uma ferramenta pedagógica encadeada no processo ensino-aprendizagem.

O conceito de avaliação formativa foi desenvolvido por Scriven (1967) e tem a função de constatar o que está sendo aprendido mediante coleta de informações feita pelo professor de forma contínua, com diversos procedimentos metodológicos e julgar o grau de aprendizagem, ora em relação a todo grupo-classe ora em relação a um determinado aluno em particular.

A avaliação formativa, na sua concepção original, não tem como objetivo classificar, selecionar, aprovar ou reprovar o aluno, mas adequar o processo de ensino ao grupo e ou àqueles alunos que apresentam dificuldades, tendo em vista o alcance dos objetivos propostos.

Nos anos vigência do Programa de Capacitação foram elaborados numerosos testes formativos para serem aplicados nos alunos nas séries e disciplinas focalizadas aplicadas nas duas séries e disciplinas. Em geral, eles apresentaram 20 questões de múltipla escolha e os alunos as respondiam no próprio caderno de teste. Após 1 hora de resolução das questões, os alunos transcreviam suas respostas para uma folha de respostas que eram recolhidas pelos professores para serem remetidas à Fundação Cesgranrio, na cidade do Rio de Janeiro, para processamento.

Para contornar o tempo necessário para a devolução dos resultados a professores e alunos, pois não estava disponível a transmissão de dados *on line*, e para que os alunos fossem imediatamente informados dos seus erros e acertos, o professor, após recolhimento das folhas de respostas, corrigia as questões com a classe, utilizando variadas formas de estimular a participação de todos os alunos. Os alunos podiam ser convidados a apresentar oralmente como chegou à resposta ou ir ao quadro para demonstrar como resolveu a questão.

A recomendação feita nas oficinas de capacitação era de que o professor aproveitasse os erros mais frequentes para reensinar os aspectos não dominados. Há bastante evidências na literatura sobre a importância do professor corrigir os erros no momento em que ele é cometido, pois a motivação é ativada quando o aluno corrige suas respostas logo após ter realizado o esforço para resolver as questões. Pesquisa realizada por Rosenshine e Stevens, em 1986, citada por Bressoux (op. cit., 2003 p.35) aponta que os professores eficazes corrigem imediatamente os erros dos alunos a fim de evitar que eles se tornem sistemáticos.

4. METODOLOGIA DO ESTUDO DOS EFEITOS DO PROGRAMA DE CAPACITAÇÃO DE PROFESSORES NO DESEMPENHO DOS ALUNOS

O objetivo desta seção é descrever o estudo feito pelos autores com 538 professores de Língua Portuguesa e de Matemática, que em 2007 estavam inscritos no Curso de Pós-Graduação lato sensu em Avaliação Escolar ministrado pela Fundação Cesgranrio e que lecionavam do 6º ao 9º ano do Ensino Fundamental da Rede Estadual do Tocantins.

Procurava-se identificar se nas escolas com professores que apresentavam um alto grau de participação nas atividades de capacitação houve aumento das médias de proficiências dos alunos, medidas no ano de 2001 pela Fundação Cesgranrio e em 2007 pela Prova Brasil. As variáveis de participação e de uso do material didático distribuído ao longo dos anos foram derivadas de dois instrumentos respondidos pelos professores e foram correlacionadas ao desempenho dos alunos mediante análises de regressão linear.

Os professores, sujeitos da pesquisa, ao inscreverem-se no curso preencheram um formulário que, além dos dados pessoais, investigava se eles já haviam participado anteriormente de atividades de capacitação promovidas pela Fundação Cesgranrio, em que ano a participação ocorreu e se aplicaram nas suas turmas a metodologia de ensino recomendada. Além dessas informações, os professores responderam, em 2008, a um questionário que retomava algumas informações sobre eles obtidas com o formulário de inscrição e incluía outras variáveis de interesse da pesquisa que, posteriormente, permitissem

correlacioná-las com o desempenho dos alunos das escolas onde eles lecionavam no período compreendido entre os anos de 2001 e 2007.

4.1. Descrição das variáveis dos professores e das escolas dos participantes do programa de capacitação consideradas no modelo de análise

O objetivo do estudo era o de procurar uma associação entre as médias de proficiência, por escola, obtidas pelos alunos das escolas estaduais do Tocantins na Prova Brasil 2007 e um grau de participação dos professores daquelas escolas nas atividades do Programa de Capacitação de 2003 até 2007, utilizando como variável controle as médias das proficiências por escola obtidas pelos alunos na Avaliação Estadual de 2001, na Escala SAEB.

4.1.1. Descrição das variáveis do professor

A ficha de inscrição continha perguntas sobre a modalidade de participação do professor nos diferentes anos do Programa de Capacitação de 2003 a 2006, inclusive, perguntas sobre o uso do material didático distribuído no período e sobre aplicação e uso dos resultados das avaliações formativas. O questionário continha a pergunta sobre o número de anos que o professor trabalhava na escola no momento de resposta ao questionário.

O primeiro passo da análise foi definir o “grau de participação” dos professores das escolas no Programa, pois nem todos os professores de 8ª série das duas disciplinas participaram do curso de Aperfeiçoamento ou do Programa de Capacitação em anos anteriores.

Uma primeira medida foi definir um índice de participação do professor a partir de 5 situações independentes: ter sido Multiplicador de 5ª a 8ª série em 2003, em 2004 e em 2005, ser participante de Seminário em 2005 e 2006.

SE VOCÊ ASSINALOU SIM, ESPECIFIQUE O ANO E A MODALIDADE DA SUA PARTICIPAÇÃO:				
MODALIDADE	ANO			
Multiplicador de 1ª a 4ª Série - EF	2003	2004	2005	
Multiplicador de 5ª a 8ª Série - EF	2003	2004	2005	
Participante do Seminário	2005	2006		

As tabelas 1 e 2 mostram o número de professores que participou como multiplicador ou nos seminários. Como se pode ver, na Matemática há mais professores que declararam ter participado nas diversas modalidades, por ocasião, do que em Língua Portuguesa. O número de professores que assistiu aos seminários é bem maior, foram cerca de 400, por disciplina.

TABELA 1. NÚMERO DE PROFESSORES POR MODALIDADE DE PARTICIPAÇÃO EM MATEMÁTICA

Matemática					
Participou	Mult03	Mult04	Mult05	Sem05	Sem06
Não	141	124	128	137	129
Sim	58	75	71	62	70

TABELA 2. NÚMERO DE PROFESSORES POR MODALIDADE DE PARTICIPAÇÃO EM LÍNGUA PORTUGUESA

Língua Portuguesa					
Participou	Mult03	Mult04	Mult05	Sem05	Sem06
Não	161	157	163	155	148
Sim	41	45	39	47	54

As variáveis do professor estão definidas no quadro 1.

QUADRO 1. DEFINIÇÃO DAS VARIÁVEIS DO PROFESSOR

Variáveis do professor	Definição das variáveis,
Grpart4	Número de participações como multiplicador ou no seminário até 2005, inclusive.
Grpart5	Número de participações como multiplicador ou nos seminários até 2006, inclusive.
Mdid	Frequência de uso de Material Didático: 2 = Muito, 1 = Pouco, 0 = Nunca.

As tabelas 3 e 4 apresentam as distribuições das duas primeiras variáveis. Nestas tabelas pode-se ver que o percentual do grau de participação de 3 ou mais, em Matemática, é mais do que o dobro do que em Língua Portuguesa.

TABELA 3. DISTRIBUIÇÃO DE GRPART4, EM MATEMÁTICA E LÍNGUA PORTUGUESA

Grau	Matemática		Língua Portuguesa	
	Número	Percentual	Número	Percentual
0	60	30.2	91	45
1	65	32.7	71	35.1
2	28	14.1	20	9.9
3	39	19.6	19	9.4
4	7	3.5	1	0.5

TABELA 4. DISTRIBUIÇÃO DE GRPART5, EM MATEMÁTICA E LÍNGUA PORTUGUESA

Grau	Matemática		Língua Portuguesa	
	Número	Percentual	Número	Percentual
0	49	24.6	76	37.6
1	40	20.1	57	28.2
2	52	26.1	44	21.8
3	45	22.6	20	9.9
4	8	4	4	2
5	5	2.5	1	0.5

A variável Mdid foi definida a partir da seguinte pergunta da ficha de inscrição.

6. SE VOCÊ RESPONDEU SIM, ESPECIFIQUE COM QUE FREQUÊNCIA USOU O MATERIAL NAS SUAS AULAS.					
Muito	<input type="checkbox"/>	Pouco	<input type="checkbox"/>	Nunca	<input type="checkbox"/>

A distribuição desta variável é apresentada na tabela 5. Cerca de 50% em cada disciplina declararam que usaram muito o material didático, entretanto observa-se um maior percentual de professores em Língua Portuguesa que nunca usou e também que usou. Na realidade, o percentual em Língua Portuguesa concentra-se nos extremos: nunca e muito.

TABELA 5. DISTRIBUIÇÃO DA FREQUÊNCIA DE USO DO MATERIAL DIDÁTICO (MDID), EM MATEMÁTICA E LÍNGUA PORTUGUESA

Frequência de uso do material	Matemática		Língua Portuguesa	
	Número	Percentual	Número	Percentual
Nunca=0	55	27.64	75	37.13
Pouco=1	47	23.62	20	9.9
Muito=2	97	48.74	107	52.97

4.1.2. Descrição das variáveis da escola

A variável da escola a ser explicada é a média de desempenho em uma disciplina – Língua Portuguesa ou Matemática – de seus alunos na Prova Brasil 2007. A variável explicativa utilizada como controle é a média de desempenho na Avaliação Estadual de 2001, na Escala SAEB.

Para as demais variáveis explicativas, procurou-se criar índices de participação da escola no Programa de Capacitação. Estas foram sempre definidas como sendo a proporção de professores na escola que satisfazem uma certa condição relacionada ao projeto, associada aos professores.

Não foi possível usar as turmas de cada professor em 2007, pois o Programa não tem dados relacionando professor e turma, além do que os controles nos anos do estudo têm que ser por escola.

O quadro 2 a seguir apresenta a descrição das variáveis da escola.

QUADRO 2. DEFINIÇÃO DAS VARIÁVEIS DA ESCOLA

Variáveis da escola	Definição
Alt_part4	número de professores da escola com $Grpart4 \geq 3$ / número de professores da escola no Curso de Aperfeiçoamento. Varia de 0 a 1.
Alt_part5	número de professores da escola com $Grpart5 \geq 3$ / número de professores da escola no Curso de Aperfeiçoamento. Varia de 0 a 1.
Alt_Mdid	número de professores da escola com $Mdid=2$ / número de professores da escola no Curso de Aperfeiçoamento. Varia de 0 a 1.

Ressalta-se que o número de professores de 5ª a 8ª séries (6º e 9º anos), por disciplina, por escola utilizado foi o total de professores (nprof) da disciplina da escola no Curso de Aperfeiçoamento. Para verificar quão adequada foi essa medida, obteve-se do Censo Educacional de 2006, o número de turmas da 5ª a 8ª série (6º e 9º anos) (ntotal) para cada escola da Rede Estadual de Tocantins. Foi feita a hipótese de que a carga horária semanal para cada uma das disciplinas de Língua Portuguesa e de Matemática é de 5 horas por semana e dessa maneira um professor com carga horária de pelo menos 30 horas pode dar aulas em pelo menos 4 turmas. Estimou-se o número de professores (nprofhat) de cada disciplina em cada escola por: arredondamento (número de turmas / 4).

A tabela 6 mostra a diferença entre o número estimado de professores por escola e o número de professores utilizado, para cada uma das disciplinas de Matemática e Língua Portuguesa e sob as hipóteses de um professor poder dar aulas para 4, 5 ou 6 turmas. Pode-se ver que mesmo no pior caso, de apenas 4 turmas, somente cerca de 30 escolas, 15% do total, têm diferenças maior que 1. Mesmo nesse caso, não foi relevante para as análises realizadas, o número de professores utilizado ou o número de professores estimado.

TABELA 6. DIFERENÇA ENTRE O NÚMERO ESTIMADO DE PROFESSORES POR ESCOLA E O NÚMERO DE PROFESSORES UTILIZADO

Nº turmas por professor	Diferença em Matemática					Diferença em Língua Portuguesa				
	0	1	2	3	4	0	1	2	3	4
4	38	64	19	9	3	44	63	14	11	2
5	73	45	12	3	0	69	51	12	2	0
6	94	32	7	0	0	94	35	5	0	0

A tabela 7 fornece a distribuição das variáveis de interesse Alt_part_4 e Alt_part_5. Vê-se que o percentual de escolas nas quais essas variáveis são maiores que ½ são de cerca de 21%, e 27%, respectivamente.

Em Língua Portuguesa, a tabela 8 mostra que esse percentual de escolas é de cerca de somente 7% para Alt_part_4 e de cerca de 15% para Alt_part_5.

TABELA 7. DISTRIBUIÇÃO DAS VARIÁVEIS ALT_PART_4 E ALT_PART_5 PARA AS ESCOLAS DO ESTUDO EM MATEMÁTICA

Proporção	Número						Percentual					
	0	0.25	0.5	0.67	0.8	1	0	0.25	0.5	0.67	0.8	1
Alt_part_4	97	1	8	2	1	25	72.4	0.7	6.0	1.5	0.7	18.7
Alt_part_5	89	1	8	2	1	33	66.4	0.7	6.0	1.5	0.7	24.6

TABELA 8. DISTRIBUIÇÃO DAS VARIÁVEIS ALT_PART_4 E ALT_PART_5 PARA AS ESCOLAS DO ESTUDO EM LÍNGUA PORTUGUESA

Proporção	Número						Percentual					
	0	0.25	0.33	0.5	0.67	1	0	0.25	0.33	0.5	0.67	1
Alt_part_4	115	0	3	7	0	9	85.8	0.0	2.2	5.2	0.0	6.7
Alt_part_5	111	1	3	8	1	10	82.8	0.7	2.2	6.0	0.7	7.5

Em relação ao uso do material, as tabelas 9 e 10 apresentam a distribuição da variável Alt_Mdid. Pode-se ver que em Matemática somente cerca de 44% das escolas têm mais da metade de seus professores que declararam usar muito. Em Língua Portuguesa, esse percentual é de aproximadamente 49%.

TABELA 9. DISTRIBUIÇÃO DA VARIÁVEL ALT_MDID PARA AS ESCOLAS DO ESTUDO EM MATEMÁTICA

Proporção	Número							Percentual						
	0	0.33	0.5	0.67	0.75	0.8	1	0	0.33	0.5	0.67	0.75	0.8	1
Alt_Mdid	61	2	11	3	1	1	55	45.5	1.5	8.2	2.2	0.7	0.7	41.0

TABELA 10 DISTRIBUIÇÃO DA VARIÁVEL ALT_MDID PARA AS ESCOLAS DO ESTUDO EM LÍNGUA PORTUGUESA

Proporção	Número						Percentual					
	0	0.33	0.5	0.67	0.75	1	0	0.33	0.5	0.67	0.75	1
Alt_Mdid	49	4	15	4	1	61	36.6	3.0	11.2	3.0	0.7	45.5

A variável do questionário utilizada foi a questão 8 (Q8) sobre o tempo de trabalho na escola, cuja resposta "D" indicava mais de 5 anos, gerando a variável dicotômica IQ8d e a "C", de 3 a 5 anos, que junto com a "D" deu origem a variável dicotômica IQ8cd (mais de 3 anos) na mesma escola.

8. HÁ QUANTOS ANOS VOCÊ TRABALHA NESTA ESCOLA?	
(A)	Há menos de 1 ano.
(B)	De 1 a 2 anos.
(C)	De 3 a 5 anos.
(D)	Mais de 5 anos.

A Prova Brasil de 2007 aplicou testes em 264 escolas urbanas da Rede Estadual do Tocantins, com pelo menos 20 alunos previstos na 8ª série (9º ano). No entanto, para a análise neste estudo, somente 130 dessas escolas têm também médias de proficiências em 2001 e professores de Matemática no Curso de Aperfeiçoamento. Em Língua Portuguesa, são 132 escolas. A variável dependente neste estudo é a média dos alunos da escola na Prova Brasil 2007 (Media07).

Na regressão descrita posteriormente, em Matemática foram identificadas duas escolas com valores aberrantes (outliers), que foram retiradas da análise.

TABELA 11. RESUMO DAS DISTRIBUIÇÕES DAS MÉDIAS DAS ESCOLAS DO ESTUDO EM 2007, 2001 E DAS VARIÁVEIS UTILIZADAS COM E SEM OS VALORES ABERRANTES, NA DISCIPLINA DE MATEMÁTICA

Matemática				
	Todas as observações		Sem as escolas com valores aberrantes	
	Média	DP	Média	DP
Media07	230,34	14,24	229,95	12,86
Media01	214,40	18,50	213,71	17,78
Alt_part_4	0,23	0,39	0,23	0,40
Alt_part_5	0,29	0,43	0,29	0,43
Alt_Mdid	0,48	0,47	0,48	0,47

TABELA 12. RESUMO DAS DISTRIBUIÇÕES DAS MÉDIAS DAS ESCOLAS DO ESTUDO EM 2007, 2001 E DAS VARIÁVEIS UTILIZADAS, NA DISCIPLINA DE LÍNGUA PORTUGUESA

Língua Portuguesa		
	Todas as observações	
	Média	DP
Media07	221,68	14,02
Media01	229,24	16,97
Alt_part_4	0,10	0,27
Alt_part_5	0,12	0,29
Alt_Mdid	0,54	0,45

TABELA 13. RESUMO DAS DISTRIBUIÇÕES DAS MÉDIAS DAS VARIÁVEIS DAS ESCOLAS COM PELO MENOS UM PROFESSOR HÁ MAIS DE 3 ANOS COM E SEM OS VALORES ABERRANTES, NA DISCIPLINA DE MATEMÁTICA

Matemática - escolas com pelo menos um professor há mais de 3 anos (Q8cd)				
	Todas as observações		Sem os valores aberrantes	
	Média	DP	Média	DP
Media07	229,56	13,24	229,76	13,16
Media01	213,90	19,27	213,43	18,81
Alt_part_4Q8cd	0,23	0,39	0,24	0,40
Alt_part_5Q8cd	0,30	0,42	0,31	0,43
Alt_MdidQ8cd	0,48	0,46	0,48	0,46

TABELA 14. RESUMO DAS DISTRIBUIÇÕES DAS MÉDIAS DAS VARIÁVEIS DAS ESCOLAS COM PELO MENOS UM PROFESSOR HÁ MAIS DE 5 ANOS COM E SEM OS VALORES ABERRANTES, NA DISCIPLINA DE MATEMÁTICA

Matemática - escolas com pelo menos um professor há mais de 5 anos (Q8d)				
	Todas as observações		Sem os valores aberrantes	
	Média	DP	Média	DP
Media07	230,52	13,57	230,84	13,44
Media01	215,10	19,43	214,41	18,79
Alt_part_4Q8d	0,32	0,43	0,32	0,43
Alt_part_5Q8d	0,41	0,46	0,42	0,46
Alt_MdidQ8d	0,58	0,46	0,57	0,46

As tabelas 11 e 12 apresentam um resumo da distribuição das proficiências em Matemática e Língua Portuguesa para as escolas do estudo na Prova Brasil 2007 e na Avaliação Estadual de 2001, assim como o resumo da diferença de médias das escolas. Pode-se ver que houve um aumento médio de cerca de 16 pontos em Matemática e um decréscimo médio de cerca de 8 pontos em Língua Portuguesa.

Estas tabelas apresentam também as médias e desvio padrão das demais variáveis explicativas. Observa-se que as médias das variáveis Alt_part_4 e Alt_part_5 são mais que o dobro em Matemática do que em Língua Portuguesa. Uma possível explicação para este fato é que um maior número de multiplicadores de Língua Portuguesa não frequentaram o curso de Aperfeiçoamento, talvez por lecionarem no Ensino Médio. Quanto à frequência do uso de material didático (Alt_Mdid), vê-se que a média é um pouco maior em Língua Portuguesa, apesar das escolas terem professor com menor participação no projeto.

As análises, em Matemática, foram repetidas duas vezes, considerando dois casos: escolas com pelo menos um professor há mais de 3 anos e escolas com pelo menos um professor há mais de 5 anos. As variáveis Alt_part_4, Alt_part_5 e Alt_Mdid foram adaptadas para esses novos casos para serem a proporção de professores nas escolas que satisfazem simultaneamente o critério da variável e do tempo na mesma escola. A essas variáveis foram acrescentados os sufixos Q8cd – mais de três anos de permanência na escola e Q8d, mais de cinco anos de permanência na escola, respectivamente para os dois casos.

As tabelas 13 e 14 apresentam as médias e desvios padrão das variáveis das escolas para esses casos.

4.2. As análises realizadas

Para verificar se há associação positiva entre o desempenho dos alunos na Prova Brasil 2007 e a participação no Programa de Capacitação conduzido pelos autores do artigo, controlado pelo desempenho dos alunos na Avaliação Estadual de 2001, na Escala SAEB, foi feita a seguinte regressão:

$$\text{Média07}_i = b_0 + b_1 * \text{Média01}_i + b_2 * \text{var}_i + \text{erro}_i$$

Onde:

Média07_i é a média dos alunos da escola i na Prova Brasil 2007,

Média01_i é a média dos alunos da escola i na Avaliação Estadual de 2001,

var_i é a variável de interesse na escola i, por exemplo Alt_part_4..

erro_i são supostas serem variáveis aleatórias independentes com distribuição normal com média 0 e mesma variância.

5. DISCUSSÃO DOS RESULTADOS

Nesta seção, inicialmente serão mostrados os resultados das regressões realizadas, respectivamente para Matemática e Língua Portuguesa. Será apresentado um estudo da magnitude dos coeficientes b₂ em relação aos desvios padrão da variável Média07 e em relação ao desvio padrão do resíduo da regressão de Média07 em relação à variável controle Média01.

5.1. Os resultados em Matemática

O primeiro modelo para Matemática utilizando a variável Alt_part_4 apresenta o seguinte resultado com o coeficiente b₂ significativo a 10% como mostra a tabela15.

TABELA 15. RESULTADOS DAS ESTIMATIVAS COM A VARIÁVEL ALT_PART_4, EM MATEMÁTICA

Coefficientes	Estimativa	Erro padrão	Valor t	Nível de significância	
b_0	136.223	11.870	11.476	0.000	
b_1	0.434	0.055	7.840	0.000	
b_2	4.718	2.605	1.811	0.073	significativo a 10%.

Após uma análise de resíduos, verificou-se que duas escolas apresentaram-se como observações aberrantes ou "outliers". Retirando-se essas escolas e refazendo-se a análise, obtém-se o coeficiente b_2 significativo a 5% como visto na tabela 16.

TABELA 16. RESULTADOS DAS ESTIMATIVAS COM A VARIÁVEL ALT_PART_4, EM MATEMÁTICA, RETIRANDO-SE DUAS ESCOLAS "OUTLIERS"

Coefficientes	Estimativa	Erro padrão	Valor t	Nível de significância	
b_0	139.755	10.985	12.722	0.000	
b_1	0.417	0.051	8.096	0.000	
b_2	4.962	2.315	2.143	0.034	significativo a 5%.

Essas escolas retiradas apresentam as seguintes características. A primeira teve uma perda de média de 257.9 para 210.8, possui 2 professores no Curso de Aperfeiçoamento, 17 turmas de 5a a 8a série pelo Censo Educacional de 2006 e supondo 1 professor para 4 turmas, precisaria ter 4 professores. Os dois professores do Aperfeiçoamento, segundo o questionário de 2008, estão há mais de 5 anos na escola, mas tiveram uma baixa participação no Programa de Capacitação tendo frequentado somente o seminário de 2005 e um deles também o seminário de 2006. A estimativa da média do Critério Brasil, segundo a pontuação de 2003, dos alunos dessa escola praticamente se manteve constante em torno de 10.5, média do limite superior da classe D, 10, e do limite inferior da classe C, 11. A taxa de aprovação no segmento da 5a a 8a série também caiu de 86.5% (txap01) para 77.6% (txap07).

A segunda escola teve um ganho grande de média, tem 1 professor no Curso de Aperfeiçoamento, mas que não respondeu ao questionário e segundo o formulário de inscrição não participou de nenhuma atividade de capacitação. Essa escola tem um nível socioeconômico mais alto, crescendo no Critério Brasil de 16.1 a 17.4, o que coloca, provavelmente, metade dos alunos na classe B2 ou acima. Sua taxa de aprovação caiu muito, de 97.6 a 83.2.

TABELA 17. CARACTERÍSTICAS DAS ESCOLAS RETIRADAS NA ANÁLISE DE REGRESSÃO EM MATEMÁTICA

Escola	Média01	Média07	Alt_part_4	Alt_part_5	nprof	ntotal	nprofhat	cb01	cb07	txap01	txap07
1	257.9	210.8	0	0	2	17	4	10.3	10.7	86.5	77.6
2	259.8	299.2	0	0	1	11	3	16.1	17.4	97.6	83.2

TABELA 18. CARACTERÍSTICAS DOS PROFESSORES DAS ESCOLAS RETIRADAS NA ANÁLISE DE REGRESSÃO EM MATEMÁTICA

Escola	Q8	IM503	IM504	IM505	IPS05	IPS06
1	D	0	0	0	1	1
1	D	0	0	0	1	0
2		0	0	0	0	0

Onde

IM503 é variável indicadora de ter sido multiplicador em 2003,

IM504 é variável indicadora de ter sido multiplicador em 2004,

IM505 é variável indicadora de ter sido multiplicador em 2005,

IPS05 é variável indicadora de ter participado do seminário em 2005,

IPS06 é variável indicadora de ter participado do seminário em 2006,

A seguir, procurou-se restringir o estudo a escolas que tinham professores na escola há mais de 3 anos. O número de professores nas mesmas escolas há mais de três anos é de 128 que estão em 96 escolas. Essas escolas podem ter também professores com menos de 3 anos.

Fez-se a mesma regressão (tabela 19) considerando somente essas escolas e a variável Alt_part_4Q8cd que é a proporção de professores há mais de 3 anos na mesma escola e com grau de participação na capacitação maior ou igual a 3. O coeficiente b_2 é significativo a 5%, quase 1%. Ressalta-se que a escola 2 não satisfaz o critério de ter professor com mais de 3 anos. A escola 1 satisfaz este critério e está entre as 96 escolas. Ela continua sendo um "outlier", e retirando-a, o resultado da análise refeita está na tabela 20, cujo coeficiente b_2 continua sendo significativo a 5%, quase a 1%.

TABELA 19. RESULTADOS DAS ESTIMATIVAS COM A VARIÁVEL ALT_PART_4Q8CD, EM MATEMÁTICA, PARA ESCOLAS COM PELO MENOS UM PROFESSOR HÁ MAIS DE 3 ANOS

Coeficientes	Estimativa	Erro padrão	Valor t	Nível de significância	
b_0	146.455	12.367	11.843	0.000	
b_1	0.381	0.058	6.608	0.000	
b_2	7.232	2.814	2.570	0.012	significativo a 5%.

TABELA 20. RESULTADOS DAS ESTIMATIVAS COM A VARIÁVEL ALT_PART_4Q8CD, EM MATEMÁTICA, PARA ESCOLAS COM PELO MENOS UM PROFESSOR HÁ MAIS DE 3 ANOS, SEM A ESCOLA 1

Coeficientes	Estimativa	Erro padrão	Valor t	Nível de significância	
b_0	136.971	11.991	11.423	0.000	
b_1	0.428	0.056	7.630	0.000	
b_2	6.573	2.663	2.468	0.016	significativo a 5%

Se o estudo for mais exigente quanto ao tempo que o professor está na mesma escola, supondo mais de 5 anos, o número de professores nesta condição é de 80 e o de escolas de 63. No entanto os coeficientes b_2 nas análises equivalentes cujos resultados estão nas tabelas 21 e 22 continuam significativos a 10%, sendo que na tabela 21 quase a 5%.

TABELA 21. RESULTADOS DAS ESTIMATIVAS COM A VARIÁVEL ALT_PART_4Q8D, EM MATEMÁTICA, PARA ESCOLAS COM PELO MENOS UM PROFESSOR HÁ MAIS DE 5 ANOS

Coeficientes	Estimativa	Erro padrão	Valor t	Nível de significância	
b_0	146.116	15.744	9.281	0.000	
b_1	0.384	0.073	5.268	0.000	
b_2	6.564	3.362	1.953	0.056	significativo a 10%.

TABELA 22. RESULTADOS DAS ESTIMATIVAS COM A VARIÁVEL ALT_PART_4Q8D, EM MATEMÁTICA, PARA ESCOLAS COM PELO MENOS UM PROFESSOR HÁ MAIS DE 5 ANOS, SEM A ESCOLA 1

Coeficientes	Estimativa	Erro padrão	Valor t	Nível de significância	
b_0	131.901	14.959	8.817	0.000	
b_1	0.454	0.070	6.526	0.000	
b_2	5.519	3.092	1.785	0.079	significativo a 10%.

O estudo continuou fazendo as análises da variável Alt_part_5 e suas derivadas. Essa variável permite a inclusão da participação dos professores no seminário de 2006. Os resultados são parecidos, mas mais “significantes”. As tabelas, a seguir, apresentam os resultados dessas análises.

TABELA 23. RESULTADOS DAS ESTIMATIVAS COM A VARIÁVEL ALT_PART_5, EM MATEMÁTICA

Coeficientes	Estimativa	Erro padrão	Valor t	Nível de significância	
b_0	136.586	11.884	11.493	0.000	
b_1	0.432	0.056	7.772	0.000	
b_2	4.292	2.390	1.796	0.075	significativo a 10%.

TABELA 24. RESULTADOS DAS ESTIMATIVAS COM A VARIÁVEL ALT_PART_5, EM MATEMÁTICA, RETIRANDO-SE SOMENTE A ESCOLA 2

Coeficientes	Estimativa	Erro padrão	Valor t	Nível de significância	
b_0	148.344	11.197	13.249	< 2e-16	
b_1	0.373	0.052	7.120	0.000	
b_2	5.243	2.208	2.375	0.019	significativo a 5%.

TABELA 25. RESULTADOS DAS ESTIMATIVAS COM A VARIÁVEL ALT_PART_5Q8CD, EM MATEMÁTICA, PARA ESCOLAS COM PELO MENOS UM PROFESSOR HÁ MAIS DE 3 ANOS

Coeficientes	Estimativa	Erro padrão	Valor t	Nível de significância	
b_0	147.951	12.323	12.006	0.000	
b_1	0.372	0.058	6.451	0.000	
b_2	7.076	2.615	2.706	0.008	significativo a 1%.

TABELA 26. RESULTADOS DAS ESTIMATIVAS COM A VARIÁVEL ALT_PART_5Q8CD, EM MATEMÁTICA, PARA ESCOLAS COM PELO MENOS UM PROFESSOR HÁ MAIS DE 3 ANOS, SEM A ESCOLA 1

Coeficientes	Estimativa	Erro padrão	Valor t	Nível de significância	
b_0	138.506	11.990	11.551	0.000	
b_1	0.419	0.056	7.443	0.000	
b_2	6.243	2.487	2.510	0.014	significativo a 5%.

TABELA 27. RESULTADOS DAS ESTIMATIVAS COM A VARIÁVEL ALT_PART_5Q8D, EM MATEMÁTICA, PARA ESCOLAS COM PELO MENOS UM PROFESSOR HÁ MAIS DE 5 ANOS

Coeficientes	Estimativa	Erro padrão	Valor t	Nível de significância	
b_0	147.505	15.468	9.536	0.000	
b_1	0.373	0.072	5.188	0.000	
b_2	7.847	3.206	2.448	0.017	significativo a 5%.

TABELA 28. RESULTADOS DAS ESTIMATIVAS COM A VARIÁVEL ALT_PART_5Q8D, EM MATEMÁTICA, PARA ESCOLAS COM PELO MENOS UM PROFESSOR HÁ MAIS DE 5 ANOS, SEM A ESCOLA 1

Coeficientes	Estimativa	Erro padrão	Valor t	Nível de significância	
b_0	133.554	14.787	9.032	0.000	
b_1	0.443	0.069	6.406	0.000	
b_2	6.464	2.975	2.173	0.034	significativo a 5%.

As análises realizadas indicam haver um efeito positivo que está de acordo com a hipótese do Programa de Capacitação de Professores ter influenciado na melhoria de desempenho dos alunos da 8ª série (9º ano) na disciplina de Matemática. Entretanto essa afirmação só pode ser feita para um conjunto pequeno de escolas que tinham professores cursando o Aperfeiçoamento e que preencheram o formulário de inscrição

e o questionário. Com esses instrumentos, não houve como avaliar a relação de outros professores e escolas nos desempenhos dos demais alunos e o Programa de Capacitação. Também não se determinou influência do Programa de Capacitação nas escolas com baixa participação declarada pelos professores do estudo, pois o Programa não atingiu a todos os professores da escola como, por exemplo, aqueles que atuam também no Ensino Médio e que também participaram das atividades de capacitação promovidas, mas que no ano de 2007 e 2008 cursavam a Especialização em Avaliação Escolar do Ensino Médio, além da rotatividade de professores nas escolas anteriormente discutida.

Nas análises de regressão feitas foram testadas outras variáveis como o uso do material didático, aplicação das formativas e discussão dos resultados com os alunos, diferença do nível socioeconômico dos alunos das escolas entre 2007 e 2001, diferença de taxas de aprovação do segmento de 5ª a 8ª (6º e 9º anos). Estas variáveis não se mostram significativas.

5.2. Resultados em Língua Portuguesa

As variáveis significativas em Matemática também foram testadas em Língua Portuguesa, mas mostraram-se não significativas. As outras variáveis mencionadas anteriormente foram também testadas e a variável Alt_Mdid, proporção de professores da escola que declararam utilizar muito o material mostrou-se ser significativa a 5%.

TABELA 29. RESULTADOS DAS ESTIMATIVAS COM A VARIÁVEL ALT_MDID, EM LÍNGUA PORTUGUESA

Coefficientes	Estimativa	Erro padrão	Valor t	Nível de significância	
b_0	141.645	14.823	9.555	0.000	
b_1	0.336	0.065	5.176	0.000	
b_2	5.373	2.427	2.214	0.029	significativo a 5%.

As análises restritas às escolas com professores com mais de 5 anos ou mais de 3 anos na escola não deram resultados significativos. Possíveis hipóteses para esse resultado é o baixo percentual de professores de Língua Portuguesa com participação em mais de 3 instâncias no Programa de Capacitação como multiplicador ou participante de seminário. Talvez, muitos multiplicadores fossem também professores do Ensino Médio, aliada à alta rotatividade de professores nas escolas pesquisadas. Outra explicação pode estar ligada às características do material didático de Língua Portuguesa que era mais detalhado e talvez mais fácil de ser assimilado e utilizado pelos professores do que o material de Matemática, que exigia a presença nos encontros com os multiplicadores nas Diretorias Regionais ou os conduzidos pelos consultores da Fundação Cesgranrio. Por outro lado o domínio dos conteúdos e da metodologia do Programa dependia nos anos de 2003 a 2005, da qualidade dos multiplicadores para que o repasse aos professores fosse efetivo.

É possível que o material didático de Língua Portuguesa, que apresentava muitos textos de diferentes gêneros, e com inúmeras sugestões de atividades de leitura e interpretação, além de orientações concretas de como utilizá-lo, tenha circulado mais amplamente nas escolas e, mesmo os professores com grau de participação pequeno nas atividades de capacitação tenham utilizado os materiais nas suas aulas, como pode ser visto na tabela onde cerca de 42% dos professores declararam ter baixa participação ($GRpart5 \leq 2$) e utilizar muito o material didático ($Mdid = 2$).

TABELA 30. NÚMERO E PERCENTUAL DE PROFESSORES DO APERFEIÇOAMENTO POR GRAU DE PARTICIPAÇÃO (GRPART5) E USO DO MATERIAL (MDID)

GRpart5	Frequência de uso do material					
	Número			Percentual		
	0	1	2	0	1	2
0	58	5	13	28.71	2.48	6.44
1	15	7	35	7.43	3.47	17.33
2	2	5	37	0.99	2.48	18.32
3	0	3	17	0	1.49	8.42
4	0	0	4	0	0	1.98
5	0	0	1	0	0	0.5

Fonte: Fundação Cesgranrio 2009

5.3. Estudo da magnitude dos coeficientes da variável de interesse

Para estudar a magnitude do coeficiente b_2 da variável de interesse, comparar-se-á este com o desvio padrão da variável dependente Média07 e também com o desvio padrão do resíduo da regressão de Média 07 em relação à variável controle Média01. Esses desvios padrão estão na tabela31 para Matemática e na tabela 32 para Língua Portuguesa. Como é de se esperar, o desvio padrão dos resíduos do modelo são menores que o desvio padrão da variável Média07. Da mesma forma, os desvios padrão são menores retirando-se as escolas com os valores aberrantes.

TABELA 31. DESVIOS PADRÃO PARA A VARIÁVEL MÉDIA07 E PARA OS RESÍDUOS DA REGRESSÃO DE MÉDIA07 EM RELAÇÃO AO CONTROLE MÉDIA01, NOS 3 CASOS EM MATEMÁTICA

	Todas as escolas	Escolas com pelo menos um professor há mais de 3 anos	Escolas com pelo menos um professor há mais de 5 anos
dp(Média07)	14.24	13.24	13.57
dp(Média07 sem valores aberrantes)	12.86	13.16	13.44
dp(resíduos)	11.68	11.02	11.40
dp(resíduos sem valores aberrantes)	10.38	10.38	10.38

TABELA 32. DESVIOS PADRÃO PARA A VARIÁVEL MÉDIA07 E PARA OS RESÍDUOS DA REGRESSÃO DE MÉDIA07 EM RELAÇÃO AO CONTROLE MÉDIA01 EM LÍNGUA PORTUGUESA

	Todas as escolas
dp(Média07)	14.02
dp(resíduos)	11.68

Observa-se na tabela 33 que a magnitude do coeficiente b_2 da variável Alt_part_4, em Matemática é considerável. Em relação ao desvio padrão da média varia de 33% deste a 55% e em relação ao desvio padrão dos resíduos de Média07 em relação a Média01 de 40% a 66%. Isto significa, que dada a mesma média M01 espera-se que a escola, na qual todos os seus professores tenham alta participação no Programa de Capacitação, tenha uma média mais alta em cerca de 50% do desvio padrão da variável Média07 em relação à escola na qual nenhum professor tenha tido participação no Programa. O comportamento do coeficiente da variável Alt_part_5, em Matemática, é semelhante, como pode ser visto na tabela 34.

TABELA 33. MAGNITUDE DOS COEFICIENTES DA VARIÁVEL ALT_PART_4 EM MATEMÁTICA

	b_2	$b_2/dp(\text{med}07)$	$b_2/dp(\text{modelo})$
Todas as observações	4.72	0.33	0.40
Sem os valores aberrantes	4.96	0.39	0.48
Escolas com pelo menos um professor há mais de 3 anos	7.23	0.55	0.66
Sem os valores aberrantes	6.57	0.50	0.63
Escolas com pelo menos um professor há mais de 5 anos	6.56	0.48	0.58
Sem os valores aberrantes	5.52	0.41	0.53

TABELA 34. MAGNITUDE DOS COEFICIENTES DA VARIÁVEL ALT_PART_5 EM MATEMÁTICA

	b_2	$b_2/dp(\text{med}07)$	$b_2/dp(\text{modelo})$
Todas as observações	4.29	0.30	0.37
Sem os valores aberrantes	4.57	0.36	0.44
Escolas com pelo menos um professor há mais de 3 anos	7.08	0.53	0.64
Sem os valores aberrantes	6.24	0.47	0.60
Escolas com pelo menos um professor há mais de 5 anos	7.85	0.58	0.69
Sem os valores aberrantes	6.46	0.48	0.62

Finalmente, a tabela 35 mostra a magnitude do coeficiente da variável Alt_Mdid em Língua Portuguesa como sendo cerca de 40% do desvio padrão da Média07 ou do desvio padrão dos resíduos da regressão da Média07 em relação a Média01.

TABELA 35. MAGNITUDE DOS COEFICIENTES DA VARIÁVEL ALT_MDID EM LÍNGUA PORTUGUESA

	b_2	$b_2/dp(\text{med}07)$	$b_2/dp(\text{modelo})$
Todas as observações	5.37	0.38	0.42

Fonte: Fundação Cesgranrio, 2009.

Para finalizar os resultados apresentados indicam que os coeficientes das variáveis de escola ligados a participação de seus professores nas atividades do programa de capacitação de Matemática são positivamente correlacionados com a variável dependente média da escola na Prova Brasil em 2007, na presença da variável controle, média da escola na avaliação de 2001.

Em Língua Portuguesa a variável que apresentou correlação positiva com o melhor desempenho dos alunos nesta disciplina expresso pela média da escola na Prova Brasil 2007 foi o uso frequente do material didático distribuído composto por apostilas e livros do professor, além de roteiros e cadernos de exercícios para os alunos.

6. CONCLUSÕES

Após seis anos de trabalho sistemático com professores e alunos do Tocantins chega-se ao final do estudo com uma indagação se valeu a pena a quantidade de recursos materiais e humanos investidos para obter-se resultados tão modestos quanto às evidências dos efeitos das ações de capacitação na melhoria de desempenho dos alunos da 8ª série do Ensino Fundamental. O estabelecimento dessas relações é de suma importância no estudo realizado, numa vez que as ações empreendidas, em todos esses anos, tiveram como principal objetivo melhorar a qualidade docente para aumentar as médias de proficiências dos alunos avaliadas pelo SAEB/Prova Brasil, entre os anos de 2001 e 2007.

Esta não é uma área de pesquisa explorada no Brasil, pois foram encontrados poucos estudos na revisão da literatura que apresentam um modelo de intervenção nos resultados acadêmicos dos estudantes de

um sistema educacional concreto, mediante propostas de mudanças da qualidade docente e o simultâneo acompanhamento do desempenho dos alunos por um período de tempo relativamente longo. No modelo do presente estudo, as médias de proficiências iniciais de uma população de alunos da 8ª série foram obtidas em 2001, os resultados de desempenho de outra população de alunos ao final de 6 anos foram avaliados em 2007 e em ambas medidas utilizou-se a mesma escala de proficiências, pretendendo-se relacionar os possíveis ganhos de aprendizagem com as ações de melhoria da qualificação docente empreendidas pelo Programa de Capacitação.

Para iniciar a discussão de resultados, recorrer-se-á novamente à Laura Goe que na revisão da literatura apresentada anteriormente (Goe, 2007, p.2) afirma que medir a qualidade docente usando os resultados de desempenho dos alunos obtidos por avaliações externas ainda é, nos dias atuais, um grande desafio por que:

- Os testes padronizados de rendimento escolar são criados para medir as proficiências dos alunos e não para medir a qualidade docente;
- É difícil separar os efeitos do professor (isto é, a contribuição líquida do professor) dos efeitos da sala de aula, tais como a relação com os pares, livros e outros materiais didáticos, clima da sala de aula, entre outros fatores;
- É difícil obter a ligação entre os dados do professor e alunos que permita conectar um professor específico ao desempenho dos alunos nos testes, a não ser em pesquisas desenhadas para tal;

Diante disto, qualquer proposta de intervenção na escola com vistas a aumentar os níveis de proficiências dos alunos tem que focalizar prioritariamente suas ações na busca da qualidade docente, na qual a qualificação dos professores é uma das componentes. Com essa afirmação não se pretende excluir outros fatores escolares, como a liderança do diretor, o clima da escola, a participação das famílias, pois eles importam e devem ser considerados, mas sim que é na sala de aula e no ambiente escolar, na interação professor-aluno que se estabelecem as relações de aprendizagem, e é nesse contexto que a qualidade docente tem que ser implementada, monitorada e avaliada.

O programa de Capacitação de Professores desenvolvido no Tocantins que foi analisado neste artigo teve como proposta, desde a sua concepção, a melhoria da qualidade docente acreditando que esta poderia ser aumentada pela qualificação dos professores e assim traduzir-se em efetividade docente, ou seja, em uma contribuição demonstrável nos desempenhos mensurados dos alunos. Desde o início, tinha-se presente que esta efetividade docente, de um lado, não é fácil de alcançar e de outro, é difícil de medir e, são essas dificuldades que serão discutidas nestas conclusões. Nelas, serão apresentadas também algumas reflexões sobre os requisitos necessários aos programas de capacitação que pretendam causar impactos na efetividade docente. Estas reflexões têm a intenção de chamar a atenção das autoridades educacionais e gestores de sistemas de ensino sobre algumas características que devem ser atendidas quando forem implementados programas de capacitação de professores com vistas à melhoria do desempenho dos alunos.

As qualificações docentes, representadas pelos diplomas, certificados de participação em cursos de educação continuada, ou mesmo aprovação em exames de ingresso na carreira docente, no entender de uma parcela de pesquisadores servem para permitir o acesso ao cargo e possibilitar aos gestores dos

sistemas educacionais, escolher entre aqueles com maior potencial para o cargo ou com probabilidade de melhor atuação em sala de aula, mas não são uma medida de efetividade docente. Para eles importa examinar quais são as relações que as qualificações docentes guardam com a efetividade docente.

No estudo realizado com os professores do Tocantins ficou provado que professores de Matemática que tinham melhorado sua qualificação participando com frequência das atividades de capacitação promovidas causaram impacto positivo no desempenho dos alunos nas escolas que lecionavam. Da mesma maneira, professores de Língua Portuguesa que, embora não tendo participado das atividades de capacitação, mas usaram frequentemente o material didático – livros e roteiros ou cadernos de exercícios para os alunos-levdados para a escola pelos multiplicadores ou por outros colegas que frequentaram as atividades de capacitação entre os anos 2003 a 2006, conseguiram aumentar a média de proficiências dos alunos nas suas escolas.

Essas evidências são também uma contribuição ao conhecimento disponível na área, uma vez que o estudo científico realizado com professores do Tocantins comprova que as atividades de capacitação de professores com certas características, conduzida por um longo período de tempo, no mínimo por 3 anos ou mais, são capazes de aumentar o desempenho dos alunos. Tais evidências permitem também afirmar que, contrariamente aos achados por Hanushek (op.cit, 2005, p.417) quanto a ser apenas uma porção pequena da variação da qualidade docente explicada pelas características observáveis de professor, como educação ou experiência, na pesquisa realizada, comprovou-se que a qualificação do professor aumentada pelas atividades de capacitação ou pelo uso do material didático e acompanhada da permanência relativamente estável na escola produziram efeitos de certa magnitude no desempenho dos alunos.

Prosseguindo com a análise de alguns autores que estudam a qualidade docente, Linda Darling-Hammond (2002) - provavelmente a mais influente pesquisadora da qualidade docente e autora de numerosas pesquisas na área-, fazendo uma revisão da literatura empírica sobre qualificações docentes e resultados dos alunos, encontra que a educação do professor é importante para as aquisições dos alunos, embora tanto a educação quanto a certificação docente possam não estar relacionadas à efetividade docente. De acordo com a autora, as pesquisas indicam que a habilidade verbal e o domínio da matéria que leciona são os mais importantes componentes da efetividade docente. Acrescenta ainda que, os professores que terminam seus cursos de formação não estão preparados para exercerem seus cargos e que algumas qualificações docentes importam menos que as outras em função do grau ou da série que o professor leciona e da matéria que ele ensina. (Darling – Hammond and Youngs, 2002 p.13-25).

Outras pesquisas examinadas no estudo e apresentado neste trabalho são unânimes em afirmar que o domínio da matéria que leciona é o primeiro requisito da efetividade docente e, este foi um aspecto central do Programa de Capacitação de Professores do Tocantins. O material didático produzido, as atividades de capacitação de multiplicadores e os seminários realizados com os professores enfatizaram o domínio dos conteúdos dos programas de ensino praticados, associado à estratégias de desenvolvimento das habilidades das matrizes de referência do SAEB.. No entanto, o Programa de Capacitação não havia realizado até 2007 qualquer atividade de monitoramento da aprendizagem dos professores. Por exemplo, no período de 2003 a 2007, os professores não foram submetidos à qualquer instrumento de avaliação que permitisse identificar lacunas dos conteúdos ensinados e pudesse gerar ações de superação de dificuldades, como o aporte de materiais didáticos complementares ou oferta de seminários extras para os professores com desempenhos muito insatisfatórios. Somente a partir do segundo semestre de 2007, com a introdução do Curso de Aperfeiçoamento em Avaliação Escolar é que foram aplicadas questões de

prova versando sobre os conteúdos ensinados, uma vez que nestes cursos, as provas e os trabalhos são requisitos necessários à concessão de títulos ou certificados. As avaliações dos docentes realizadas nos anos 2007 e 2008 revelaram que quantidade expressiva dos professores de Matemática não dominaram os conteúdos ensinados, pois 44% deles (118) foram reprovados no final do curso de Aperfeiçoamento. Em Língua Portuguesa a reprovação foi menor, 16% dos 278 professores. Estas reprovações revelam que, pelo menos em Matemática, o primeiro requisito para a efetividade docente não foi atendido integralmente, embora nas regressões realizadas, a aprovação no Curso de Aperfeiçoamento não apresentou correlação positiva com o desempenho dos alunos.

É interessante lembrar que a maioria das pesquisas sobre certificação docente encontra uma correlação positiva entre os resultados do desempenho dos professores nos testes de Matemática (Goldhaber and Brewer, 1999; Harris and Sass, 2007, Hill Rowan and Ball, 2005) e o desempenho dos estudantes, particularmente no 2º segmento do Ensino Fundamental. Entretanto, no caso dos professores do Tocantins, uma explicação provável para a não correlação positiva pode ser o descompasso de tempo entre a medida de desempenho dos alunos realizada pela prova Brasil em 2007 e a avaliação final dos professores ocorrida em novembro de 2008. Seria possível supor, por exemplo, que os professores adquiriram ou melhoraram o domínio dos conteúdos lecionados com as atividades e os materiais didáticos distribuídos para eles e suas turmas ao longo do curso, iniciado apenas em outubro de 2007 e continuado até o final de 2008. Acrescente-se também que não foi feito qualquer processo inicial de seleção dos participantes até 2007, uma vez que nos anos anteriores competiu à Secretaria de Educação do Tocantins indicar os professores entre aqueles que estavam em exercício na Rede Estadual, mesmo os que não atendessem aos requisitos para a função, como por exemplo, não possuísem licenciatura plena nas duas áreas curriculares do 2º segmento do Ensino Fundamental. Mais uma vez, somente com a oferta do Curso de Aperfeiçoamento, pode-se ter o controle da titulação inicial dos professores.

Na literatura examinada constatou-se também que os professores efetivos, além dos conteúdos do programa, devem dominar os métodos para ensiná-lo, especialmente em Matemática e Ciências (Walsh and Tracy, 2004 e Allen, 2003), áreas onde há sólidas evidências sobre o impacto do domínio do conhecimento e da pedagogia utilizada pelos professores nos resultados dos estudantes, embora estes devam ser requisitos presentes também nos professores de todas as disciplinas do currículo (Darling-Hammond, 2008).

No que poder-se-ia chamar de "boa pedagogia" as pesquisas indicam que o engajamento dos estudantes nos processos de aprendizagem é a chave de um ensino efetivo. Neste sentido o Programa de Capacitação propôs a criação de situações de trabalho em grupo com os alunos, o incentivo da aprendizagem cooperativa, a distribuição de roteiros e cadernos de exercícios para as turmas. Estas estratégias didáticas tinham como objetivo modificar a prática docente diminuindo a frequência das atividades de ensino frontal e aumentando as situações didáticas mais dinâmicas com os alunos. Este foi, sem dúvida, um componente importante do Programa de Capacitação que não foi avaliado, já que não foi possível contratar monitores/supervisores, externos à Secretaria de Educação, para realizar as atividades de acompanhamento dos professores nas escolas. Os multiplicadores que intermediavam as ações de capacitação junto aos professores entre os anos de 2003 a 2006, não chegaram até as salas de aula. Acrescente-se o fato de muitos multiplicadores não possuírem também os requisitos necessários para exercerem as funções de monitoramento, pois eram profissionais do quadro docente, que, embora

interessados e disponíveis, necessitavam também de maior suporte técnico, logístico e de supervisão por parte do Programa.

Entretanto o monitoramento externo ou realizado por professores mais experientes da própria escola são aspectos considerados prioritários pelos pesquisadores que estudam a melhoria da qualidade e da efetividade docente. Darling-Hammond e demais autores do texto *Democracy at Risk* (2008, p.29-30) propõem a criação de um programa Federal nos Estados Unidos para todos os professores iniciantes na carreira, garantindo a presença de um profissional altamente qualificado e especializado em treinamento docente na escola, ou em um conjunto de escolas geograficamente próximas, para dar suporte técnico e pedagógico aos professores na sua prática cotidiana.

Os autores sugerem também a criação de programas de residência de professores (por analogia aos programas de residência médica) para as comunidades socialmente mais vulneráveis, realizados em parceria com as universidades. Estes programas seriam mantidos pelo governo e os candidatos às residências, de no mínimo um ano, seriam pagos durante o período de formação e em troca assumiriam o compromisso de permanecerem na localidade a eles destinada por 4 ou 5 anos. A permanência dos professores por um período mais longo em uma mesma escola é um fator que deve ser controlado pelos implementadores de programas de melhoria da efetividade docente.

Retomando o aspecto da ausência de controle da aprendizagem e da prática em sala de aula dos professores, o Programa de Capacitação, também não tinha como empreender qualquer ação efetiva para evitar a rotatividade dos professores. Este é um aspecto que não deve ser minimizado, pois sabe-se que a mobilidade dos docentes, tanto dentro da rede quanto para fora dela é um fenômeno de uma certa magnitude, principalmente quando os sistemas educacionais mantêm um quadro expressivo de professores temporários.

Na presente pesquisa, a rotatividade dos professores no Programa de Capacitação só pode ser parcialmente dimensionada, uma vez que os dados disponíveis foram obtidos com os professores inscritos no Curso de Aperfeiçoamento, que ao preencherem o formulário de inscrição assinalaram quanto tempo e qual a modalidade de participação nas atividades promovidas nos anos anteriores. Não há informação sobre demais professores que, embora possam ter participado das atividades anteriormente, não puderam ser alcançados pela pesquisa.

A variável participação definida nas páginas anteriores deste artigo mostrou ter um efeito positivo com o desempenho dos alunos em Matemática, com nível de significância a 1% quando o professor está há mais de 3 anos na escola e a proporção de professores na escola com grau de participação 3 ou mais nas 5 modalidades de participação, ou seja, ter sido multiplicador ou participante dos seminários.

Uma provável explicação também para a pouca permanência dos professores pesquisados nas escolas pode ser a situação trabalhista de vínculo precário. Dos respondentes, 27% dos de Língua Portuguesa eram prestadores de serviço por contrato temporário e em Matemática esse percentual era de 31%. Apenas um pouco mais de 50% das duas disciplinas afirmaram ter um vínculo permanente na carreira docente.

Um outro aspecto central do Programa de Capacitação voltado para o acompanhamento sistemático do desempenho dos alunos mediante a aplicação de avaliações formativas, na prática, foi muito difícil implementá-lo com eficiência. As questões de natureza logística de impressão, distribuição, recolhimento, processamento de posterior análise de resultados impossibilitaram que as avaliações formativas

cumprissem um dos seus papéis mais relevantes de acompanhamento do desempenho do aluno, que é o de fornecer informações sistematizadas sobre esses resultados para todos os níveis do sistema de ensino: turmas, escolas, e Secretaria. Assim, da mesma maneira que o Programa de Capacitação não foi capaz de prover o monitoramento das atividades e do uso do material didático pelos professores na sala de aula também não pôde controlar se os professores corrigiam as questões das avaliações formativas imediatamente após sua aplicação e, quando necessário, organizassem atividades de reforço. Essa prática de fornecer aos alunos um feedback imediato sobre seus erros e acertos é apontada, em quase todas as pesquisas dos efeitos de fatores escolares sobre o desempenho dos alunos, como uma condição para o ensino efetivo como apresentado em algumas seções deste texto.

Estes e outros aspectos discutidos podem dar uma dimensão das dificuldades de atuação em sistemas educacionais concretos, onde algumas variáveis intervenientes nas propostas de melhoria não são passíveis de controle pelos agentes promotores das ações e assim, reduzem as possibilidades de alcançar metas projetadas. Entretanto, são bastante estudadas as dificuldades enfrentadas pelos processos de promoção de mudanças que envolvam modificações de práticas arraigadas, atitudes, crenças e valores, especialmente quando estão envolvidos um grande número de indivíduos dispersos em uma extensão territorial considerável.

Ignorar essas premissas é adotar uma postura de que o professor não faz melhor o seu ofício apenas porque não sabe fazê-lo e não, entre outros determinantes, porque não pode ou porque não quer. O desafio nos programas de capacitação de professores que, precisa incluir um expressivo número de docentes e escolas, e lidar com as múltiplas variáveis dos sistemas educacionais concretos está em superar a aparente oposição entre a dimensão individual, biográfica dos professores para articulá-la a um contexto mais amplo de projetos de melhoria da qualidade e equidade da educação oferecida a todos alunos.

Refletindo sobre esses mais de 6 anos de experiência com os professores do Tocantins é inevitável, no final, avaliar os erros e acertos, repensá-la para retirar dela elementos que possam servir, não de conselhos, mas de contribuição para um debate sobre programas de melhoria da qualidade docente que pretendam aumentar o desempenho dos alunos.

O primeiro aspecto a ser discutido é o que diz respeito à ênfase que deve ser dada pelos programas de capacitação docente ao domínio da matéria, investindo fortemente no ensino dos conteúdos dos programas das disciplinas da Educação Básica e garantindo que os professores efetivamente os aprendam. Para tal, os programas devem ser relativamente longos, sistemáticos e apresentar mecanismos de avaliação frequentes dos professores, aplicando testes, exercícios e ou trabalhos sobre os conteúdos ensinados. É importante que os programas de capacitação criem também as condições de reforço para os professores que não alcancem os níveis adequados de aprendizagem, já que a reprovação não é desejável.

Da mesma forma, é importante ensinar o professor a transpor os conteúdos aprendidos para seu dia a dia na sala de aula, criando situações dinâmicas de aprendizagem e para isso, os programas de melhoria da qualidade docente devem desenvolver materiais de formatos variados, como roteiros, cadernos de exercícios para o aluno, Dvds com simulações de situações concretas e como o professor deve atuar na sala e na escola. Esses Dvds, além de demonstrarem o ensino de aspectos críticos do ensino dos conteúdos, podem também abordar questões sobre indisciplina, falta de motivação e violência, dinâmica

do trabalho em grupos cooperativos, entre outros. É possível também, nos dias atuais usar o computador para manter o professor em uma rede de comunicação entre os consultores e coordenadores das capacitações e com os demais colegas do curso. O acompanhamento do desempenho dos alunos deve ser realizado frequentemente e deve ser fornecido um banco de questões de avaliação formativa para serem aplicadas nos alunos e colocá-lo à disposição das escolas e dos professores. É importante por outro lado, ajudar ao professor a interpretar os resultados obtidos e os programas de capacitação têm que ensinar aos professores a dominar a tecnologia de avaliação e de interpretação de escalas de proficiências, empregada com sucesso na última década nas avaliações externas.

O acompanhamento do desempenho dos alunos não pode, por outro lado, dispensar medidas de aprendizagem externas e assim os programas de capacitação precisam criar ou aproveitar mecanismos de avaliação externos às escolas.

Por fim, sugere-se o debate sobre três aspectos importantes que precisam ser desenvolvidos e divulgados.

O primeiro trata dos mecanismos de monitoramento dos programas de capacitação dos professores na sala de aula e nas escolas, que tenham características de controle da efetividade docente e de apoio para as dificuldades encontradas. A formação a supervisão e o acompanhamento desse profissional altamente especializado é um desafio que precisa ser enfrentado em conjunto com as Universidades, o Ministério de Educação e Secretarias.

Uma outra sugestão para o debate é o estudo da criação de Programas de Residência Docentel, feitos em parcerias com as Universidades e que poderiam transformar a médio prazo os Cursos de Licenciaturas. A última questão levantada para o debate diz respeito à carreira docente sugerindo que, os estudos que estão hoje nas agendas dos Ministérios de Educação incluam também na parcela variável da remuneração docente, além da titulação e anos de serviço, o desempenho dos alunos.

Para concluir retomam-se as recomendações do Relatório McKinsey que estudou os melhores sistemas educacionais do mundo, apresentado na Introdução deste texto e no qual aqueles que estavam situados nas primeiras posições apontaram que para garantir altos desempenhos dos alunos deve-se:

- atrair as pessoas certas para a carreira docente;
- transformá-las em instrutores efetivos;
- garantir que a escola e o sistema sejam capazes de fornecer a melhor instrução possível para cada aluno.

REFERENCIAS BIBLIOGRÁFICAS

- Allen, M. (2003). *Eighth questions on teacher preparation*. Denver, CO: Education Commission of the States.
- Alliance for Excellent Education, (2008). *Issue Brief*. March. Washington, D.C: Met Life Foundation.
- Braun, H. (2004). *Value-Added modeling: what does due diligence require?*. Princeton, N.J: Educational Testing Service.
- Bressoux, P. (2003). As Pesquisas sobre o Efeito Escola e o Efeito Professor. Educação em Revista, *Revista da Faculdade de Educação da UFMG*, n38, dez, Belo Horizonte.

- Clotfelter, C., Ladd, H., Vigdor, J. (2007). *Teacher credentials and student achievement in high school: a cross-subject analysis with fixed effects*. Washington, DC: Urban Institute.
- Darling – Hammond, L. (2000). Teacher quality and student achievement: A review of state policy evidence. *Education Policy Analysis Archives*, 8(1).
- Darling – Hammond, L., and Youngs, P. (2002). Defining “highly qualified teachers”: What does “scientifically –based research” tell us?. *Educational Researcher*, 31(9), pp. 13-25.
- Darling – Hammond, L., and Bransford, J. (2005). *Preparing Teachers for Changing World: what teachers should learn and be able to do*. San Francisco: Jossey-bass.
- Goe, L. (2007). *The link between teacher quality and students outcomes: a Research Synthesis*. Washington, DC: National Comprehensive Center for Teacher Quality.
- Hanushek, E.A. (1971). Teacher Characteristics and Gains in Student Achievement Estimation: using micro data. *American Economic Review*, 61(2), pp. 280-288.
- Hanushek, E.A., Kain, J.F. Rivkin, S.G. (2005). Teachers, Schools and Academic Achievement. *Econometrica*, 73(2), pp. 417-458.
- Harris, D.N. and Sass, T.R. (2007). *Teacher training, teacher quality and student achievement*. Washington, DC: National Center for Analysis of Longitudinal Data in Education and Research.
- Hill, H.C., Rowan, B.E., Ball, D.L. (2005). Effects of teacher mathematical knowledge for teaching on student achievement. *American Educational Journal*, 42(2).
- Noell, G.H. (2006). *Annual Report of: Value added assessment of teacher preparation*. Baton Rouge: Louisiana State University.
- Reynolds, D., Hopkins, D., et Stoli, L. (1993). Linking School Effectiveness Knowledge and School Improvement Practice: towards a synergy. *School Effectiveness and School Improvement*, 4, pp. 37-58.
- Rivikin, S., Hanushek, J. (2002). *Teachers, Schools and Academic Achievement*. Texas Schools Project. Dallas: University of Texas-Dallas.
- Rosenshine, B., et Stevens, R. (1986). Teaching Functions, in WITTRICK, M.C. (ed.). *Handbook of Research on Teaching*. 3ª édition, pp. 376-391. New York: Macmillan.
- Scriven, M. (1967). The Methodology of Evaluation. in Tyler, R.W., Gagné, R.M. and Scriven, M (Eds), *Perspectives of Curriculum Evaluation*. pp.39-83, Chicago: Rand McNally.
- The Forum Education Democracy, (2008). *Democracy at Risk: the need for a new federal police in education*. Washington, DC.
- Slavin, R.E., Hurley.E.A. and Chamberlain, A.M. (2003). Cooperative Learning and Achievement :Theory and Researchs in Reinolds W.M. and Miller, G.E. (eds). *Handbook of Psychology*, 7, pp. 177-198, NJ: Hoboken
- Veldman, D.J et Brophyj, E. (1974). Measuring Teacher Effects on Pupil Achievement. *Journal of Educational Psychology*, 66(3), pp. 319-324.
- Walsh, K and Tracy, C. (2004). *Increased the odds*. Washington, D.C: National Council on Teacher Quality.

CUATRO HIPÓTESIS SOBRE LA INSTITUCIONALIZACIÓN DE LA EVALUACIÓN EDUCATIVA EN AMÉRICA LATINA: EL URUGUAY EN MIRADA COMPARADA

FOUR HYPOTHESES ON THE INSTITUTIONALIZATION OF EDUCATIVE
EVALUATION IN LATIN-AMERICA: URUGUAY IN COMPARATIVE
PERSPECTIVE

Tabaré Fernández

Revista Iberoamericana de Evaluación Educativa 2010 - Volumen 3, Número 3

<http://www.rinace.net/riee/numeros/vol3-num3/art5.pdf>

Fecha de recepción: 31 agosto de 2010
Fecha de dictaminación: 24 octubre de 2010
Fecha de aceptación: 03 de noviembre de 2010

¿Por qué se observa una heterogeneidad tan importante en la institucionalización de la evaluación de aprendizajes en la educación básica y media de América Latina? Actualmente, sólo existe un reducido grupo de países donde *regularmente* se ha evaluado el aprendizaje *en varios grados*, con metodologías que permiten *comparar* los resultados a lo largo del tiempo y *el público accede* a los informes. La norma general es que alguno o varios de estos atributos han estado ausentes. Pasados casi treinta años y contando con una significativa cantidad de evaluaciones en la región, son varios los países en los que hubo programas ya no tienen, directamente no han podido establecer ninguno.

Uruguay es un caso destacado entre éstos últimos. Inició en 1996 las evaluaciones de aprendizaje periódicas, estableciendo un programa de evaluación en 6tos. Primaria con amplia difusión, contextualización y análisis de los resultados al punto que fue citado como un modelo en la región (Benveniste, 2000 y 2002; Ravela *et al.*, 2000). En el año 2000, se estimaba que el 44.8% de los maestros de Primaria habían leído al menos uno de los informes de evaluación (Fernández *et al.* 2000: cuadro IV.4), guarismo que se mantenía en el 45.8% para el año 2003 (Fernández and Midaglia, 2005: cuadro 4.1), con un menor impacto sobre la planificación de la enseñanza de aula y una valoración positiva de los instrumentos (Aristimuño and Kaztman, 2005: cuadro 7). Una década más tarde, la Administración Nacional de Educación Pública (ANEP)¹ no ha logrado equiparar las últimas pruebas aplicadas en 6tos. Primaria y no ha logrado establecer ningún programa para evaluar la educación media, siendo PISA el único mecanismo para monitorear la calidad en este nivel. A esto cabe añadir episodios calificables de censura de resultados, el primero ocurrido en 2003 y el último relacionado con las evaluaciones que la Secundaria Técnica Pública viene aplicando a sus estudiantes desde 2007. La Ley General de Educación (n°18437) promulgada en diciembre de 2008, procuró abrir una nueva etapa en este ámbito de las políticas educativas creando el Instituto Nacional de Evaluación Educativa (INEE) como entidad para-estatal, independiente de la ANEP, pero hasta el presente (2010), el Poder Ejecutivo aún no ha dado pasos concretos para establecer el INEE.

En otros trabajos, he explorado aspectos históricos y organizacionales *singulares* que refieren a este proceso (Fernández *et al.*, 2000; Fernández and Midaglia, 2003 y 2005; Fernández, 2008 y 2009). Aquí deseo explorar hipótesis explicativas que relacionan los niveles de institucionalización de la evaluación en los países de la región, con las reformas de segunda generación (en particular, la descentralización), la orientaciones ideológicas de los gobiernos (el peso del mercado y del ciudadano), utilidad técnico-pedagógica la información y el diseño institucional del gobierno de las agencias de evaluación. Este trabajo me permitirá discutir para el caso específico, el problema politológico más general de la formación de instituciones en contextos de los conflictos redistributivos experimentados por la educación básica y media desde los años noventa; conflictos de los que Uruguay por cierto no ha escapado.

1. LAS HIPÓTESIS

1.1. Hipótesis sobre la institucionalización

¿Qué factores pueden explicar la débil institucionalización que tendría hoy en día la evaluación educativa

¹ En Uruguay, la máxima autoridad educativa no es el Ministerio sino este cuerpo colegiado de competencia nacional y de organización hipercentralizada, que administra los niveles inicial, primaria, secundaria general, secundaria técnica y formación docente inicial.

en Uruguay, a pesar de sus vigorosos inicios? En la bibliografía sobre la evaluación educativa de América Latina se pueden rastrear hasta cuatro grandes teorías que permiten construir explicaciones rivales, aunque no contrapuestas, sobre la formación de instituciones de evaluación. Si bien no son necesariamente contradictorias, cada una pretende para sí el papel de explicación principal.

La teoría más recurrida, al menos en los años noventa, asociaba la existencia y pervivencia de las evaluaciones a la puesta en marcha de reformas de segunda generación en el sistema educativo (DiGropello, 1999, 2004; Winkler and Gershberg, 2000), y a las condiciones puestas por recibir préstamos de parte de los organismos multilaterales tales como el Banco Mundial y el Banco Interamericano (Narodowski, Nores and Andrada, 2002; Bertolini, 2001). En particular, la descentralización de la prestación de los servicios a unidades subnacionales, conllevó como contrapartida, nuevos roles para “ministerios de educación sin escuelas”, fundamentalmente relacionados al control de desempeño de escuelas a las que se le atribuían crecientes competencias (Fernández, 2004).

Los procedimientos estándares de generación de información estadística y pedagógica dentro de los sistemas, más allá de cuál fuera el grado de desarrollo alcanzado, resultaron insuficientes para que el ministerio evaluara la *calidad* de la educación; fue necesario un *sistema de evaluación* y así quedó consagrado en las leyes de educación de ese período. Las evaluaciones externas, legitimadas ampliamente en el diseño de políticas educativas en Estados Unidos, resultaron ser el mecanismo ideal (Narodowski, Nores and Andrada, 2002). En consecuencia de este argumento *sistémico* o *estructural*, debiera observarse que las evaluaciones están más sólidamente institucionalizadas en aquellos países donde más se avanzó en reformas de segunda generación, sea por la vía de la descentralización o la privatización (hipótesis 1). El indicador más general será si el ministerio es el principal prestador de servicios de educación básica.

La segunda teoría vincula la fortaleza de la institucionalización de la evaluación a la *orientación liberal* de un gobierno. En un mercado signado por una fuerte asimetría entre prestadores y clientes, dichas políticas asignarían un papel crucial a la información en la toma de decisiones de las familias. Claramente, esto modificaría los comportamientos agregados al declinamiento de la calidad de la educación, y demandaría la habilitación de mecanismos de salida o de voz (Hirshman, 1970). Esta visión sobre el funcionamiento de la prestación educativa genera un conflicto dado que propone *empoderar* a los actores externos a la escuela (incluido aquí la jerarquía ministerial) y en general *a los ciudadanos*, frente a los maestros y directivos de las escuelas. Aquellos serían los actores más interesados en conservar las evaluaciones porque la institucionalidad educativa disminuye a su favor las asimetrías de información. Dentro de esta visión *liberal*, la difusión de resultados tanto a nivel *general* como *a nivel de cada escuela*, es una propiedad crítica de cualquier evaluación. De aquí el impulso por establecer evaluaciones con dos características principales: periodicidad (anual por ejemplo) y cobertura censal (para poder entregar información de cada escuela).

Este argumento permitiría sustentar dos hipótesis sutilmente distintas. La primera hipótesis asociaría la estabilidad de partidos liberales con una mayor institucionalización de la evaluación, y la alternancia “derecha /izquierda” con la supresión los programas de evaluación (hipótesis 2b). Aunque así planteada puede parecer muy restringida, tiene una clara conexión con una de las teorías generales que explican la formación de instituciones a través de la *estabilidad* en el tiempo de las reglas, de la inercia, y de los costos implicados en el cambio de reglas que ya han estabilizado expectativas sociales (North, 1995). La *estabilidad* ha sido un valor particularmente importante para el liberalismo político en la región dado que

en su visión, esto permite a todos los actores (públicos y privados) ajustar expectativas recíprocas, costos y beneficios, partiendo de la base de los resultados anteriores y sin emplear tiempo y recursos adicionales para acceder a información adicional. La segunda hipótesis es más sutil y considera exclusivamente la orientación de la política de difusión de resultados que desarrolló la agencia de evaluación frente a cada proyecto. En el límite, si hubo libre acceso o si por el contrario, se introdujo algún tipo de censura a la comunicación de resultados al público (en adelante, hipótesis 2b).

En este trabajo se testearán ambas hipótesis, definiendo cuatro indicadores: i) el número de evaluaciones realizadas en cada país; ii) el total de años cubiertos entre la primera y la última evaluación; iii) y si actualmente hay evaluaciones ("vigencia"); y iv) si hubo evaluaciones en las que no se difundieron resultados al público².

La tercera teoría define la formación de las instituciones de evaluación con base en su *utilidad* (pedagógica) para actores educativos (maximizadores): maestros, inspectores, autoridades. Supondré que toda evaluación bien diseñada es un instrumento óptimo para concretar los intereses pedagógicos que motivan a los actores³. El argumento deposita especial atención a las distintas dimensiones de la hechura *técnica* de la política y al seguimiento de *buenas prácticas* en materia de evaluación: clara definición de los conocimientos a evaluar para cada área y grado ("tabla de especificaciones"); sólida validez de constructo; coherente aplicación de una teoría de la medición (clásica o TRI); diseño muestral o censal tal que les permite realizar inferencias válidas para la población estudiantil delimitada; equiparación de las pruebas para comparar en el tiempo; contextualización de resultados con indicadores sociales; análisis multivariados de factores asociados; comunicación a tiempo de resultados a las escuelas; redacción de informes amistosos y elocuentes, acompañados de recomendaciones didácticas (Ravela, 2001; Ravela *et al.*, 2008). Ahora bien, si todas las evaluaciones tienen altos costos, las buenas lo son aún más (tanto para los maestros como para los ministerios). Pero un buen diseño, que obtenga un consenso sustantivo de los actores en materia pedagógica, produciría beneficios que excederían ampliamente a los costos públicos y privados. Estos extremos son más complejos de evaluar y aquí no puede ser discutido en toda su extensión, aunque hay estudios de caso que profundizaron en estos aspectos (Pinkasz *et al.*, 2005; Fernández and Midaglia, 2005; Aristimuño and Kaztman, 2005; Zorrilla, 2003; Narodowski, Nores and Andrada, 2002). En términos comparados, contrastar esta teoría requeriría de máxima, evaluar directamente las evaluaciones realizadas, un trabajo de extensión que sobrepasa los recursos disponibles. En su lugar, aquí adoptaré una hipótesis mínima que reportará como indicador dicotómico si existen cambios declarados por las propias agencias en la metodología empleada para hacer evaluación. Esto permitirá sondear una de las dimensiones consideradas más atrás: la *comparabilidad en el tiempo de los resultados generada por una metodología consistente* (hipótesis 3).

1.2. Agente principal en la evaluación educativa

La bibliografía neo-institucionalista sobre educación ha resaltado constantemente que la educación es un caso típico de asimetría de información en la que, el cliente principal de la relación educativa (el alumno y sus familias), carece de información válida sobre el resultado final del proceso educativo: ¿qué tipo de

² La duración de este período debiera definirse laxamente, puesto que depende en primer lugar de las capacidades técnicas existentes en las agencias de evaluación, pero también, depende de la presión oficial por conocer u ocultar resultados. En este balance, el período de un año luego de aplicada una prueba es un tiempo más que razonable para definir censura.

³ Refiero tanto a las externas como internas. El supuesto es fuerte pero encuentra total consenso en todas las teorías pedagógicas contemporáneas.

aprendizajes ha hecho el alumno y qué nivel de competencia ha alcanzado? (Chubb and Moe, 1990; Gauri and Wawda, 2003)⁴.

Las calificaciones de las maestras, si bien tienen valor jurídico (implican por ejemplo, la promoción o reprobación de un grado), carecen de validez informativa clara: ¿una calificación de 6 en esta escuela del sur es comparable a la asignada en una escuela al norte del país?. En la Educación Media Superior por lo general esta falta de validez suele afectar las trayectorias académicas en forma significativa cuando los estudiantes, que emigran de una localidad a una capital, encuentran que sus desempeños han sido sobre-valorados y que sus posibilidades de progresar son mucho menores que las imaginadas.

Este argumento clásico también se puede extender y se puede agregar al conjunto de la ciudadanía para hacer referencia a la asimetría de información que existe entre la autoridad representativa (por ejemplo, el Parlamento) y el sistema educativo. En varios países, los legisladores que deciden anualmente sobre montos presupuestales para la educación, tienen información escasa: a lo sumo sobre la tasa de reprobación y la tasa de abandono de cursos. También es razonable suponer que en la misma situación de asimetría se encuentran otras autoridades del Poder Ejecutivo, tales el Presidente y ministerios de otras ramas no educativas, incluso cuando tienen directo interés en los resultados educativos o realizan transferencias importantes para financiar proyectos en la educación.

Los estudios sobre uso de la información para la formulación de políticas educativas hechos en la región (Cueto, 2005), muestran que los ministerios o secretarías de educación también disponen una reducida y endeble información válida sobre el aprendizaje de sus alumnos. Tampoco cuentan en muchos casos, con información sobre variables clave de los procesos: tal como asistencia docente, la implementación de proyectos educativos institucionales o la realización de coordinaciones docentes. Se puede inferir por tanto, que esta relación entre los docentes y las autoridades se desarrolla en una asimetría de información tal, que hace imposible controlar las condiciones de desempeño de las escuelas.

La introducción de una evaluación externa, estandarizada, periódica, con base en un diseño metodológico documentado, con una teoría de la medición y un muestreo rigurosos, acompañada de cuestionarios de contexto, asegura indicadores válidos y confiables. Las escuelas dejan de ser "cajas negras". Es claro entonces, que la evaluación tiene potencialidades para alterar la asimetría de información de la relación educativa (al menos) en favor de los tres principales identificados: familia, parlamento y las propias autoridades educativas⁵. Algunas de las consecuencias de política que acarrea esa nueva posición informada del principal pueden ser perjudiciales para los docentes: atar salarios a desempeños (Carrera Magisterial de México o Chile); asignar partidas presupuestales a metas de aprendizaje (ENIA en Uruguay); focalizar programas socio-educativos en escuelas de bajos desempeños (el P900 de la Concertación en Chile entre 1990 y 1994); etc.

Ahora bien, a esta altura del desarrollo de mi argumento se hace necesario introducir el segundo componente del marco teórico: la teoría de la agencia. Junto con examinar cómo la evaluación modifica

⁴ La bibliografía también destacada otras características de esta relación: i) la dificultad de medir estos resultados educativos; ii) la dificultad de programar las tareas propias del proceso de aprendizaje; iii) la falta de una "tecnología" (blanda) estándar para el desarrollo de las tareas. Estos aspectos generan altas incertidumbres sobre el desarrollo y resultado de la tarea.

⁵ Los neo-institucionalistas no suelen tratar el tema sobre el aislamiento didáctico y pedagógico en que suelen estar los docentes, más fiscalizados que asesorados, y con reducidas oportunidades de comparar *formativamente* sus prácticas y sus resultados con otros colegas. Este aspecto generalmente ignorado, ha sido largamente trabajado en el contexto del Grupo de Trabajo sobre Estándares y Evaluación; véase por ejemplo, Ravela *et al* (2008) y (2009).

la relación educativa, es necesario examinar sobretodo, qué relación se establece entre la agencia de evaluación y la autoridad ministerial; el primer principal de esta relación evaluativa. En consecuencia, propondré que el estudio de las evaluaciones educativas desde el punto de vista institucional, debiera concentrarse en el examen de las modificaciones establecidas en dos aspectos importantes conectados con la modificación de la asimetría principal/agente: i) la producción de información *periódica* sobre los aprendizajes, y ii) el acceso a la información de parte de la pluralidad de principales definidos más arriba⁶.

Además de los supuestos clásicos que la teoría hace sobre el agente (Eisenhardt, 1989), supondré que: i) las agencias de evaluación están interesadas en hacer más y mejores evaluaciones, ya que estos dos son elementos que les aportan prestigio en el campo internacional de los evaluadores; y ii) su aversión al riesgo se reduce a evitar fracasos en el desarrollo del operativo de la evaluación (por ejemplo, por falta de apoyos políticos o por bloqueos sindicales). Sin detallar la importancia diferencial que puedan tener estos dos objetos de comportamiento a regular, sostendré que la institucionalidad de la evaluación educativa es débil cuando el examen de los casos nacionales concluye señalado la *carencia* de alguno de aquellos elementos.

La hipótesis 4 que propongo reconoce la eficacia causal que sobre la formación institucional tienen las reformas de segunda generación (hipótesis 1), el tiempo y de las ideologías, (hipótesis 2a), la importancia difusión pública de resultados (hipótesis 2b) y la consistencia técnico-pedagógica y metodológica (hipótesis 3). Sin embargo, considera que este tipo de análisis desconoce la naturaleza misma de toda institución y de los conflictos de intereses específicos que existen entre evaluadores y ministerios. De aquí que sostengo que *la mayor o menor institucionalización de la evaluación es consecuencia del diseño institucional que regula, prescribe, habilita, las relaciones entre la agencia de evaluación y el primer principal de la relación evaluativa*. Los obstáculos mayores radican en que el primer principal de la relación evaluativa, el ministerio, no siempre comparte la opinión de que es necesario informar o que debe seguirse evaluando. Se trata de un actor que también tiene aversión al riesgo, en este caso, a la exposición pública de las ineficiencias y de la mala calidad⁷. Razón por lo cual, restringe o prohíbe la difusión, modifican la ubicación de las agencias en la estructura ministerial, reducen los salarios, desjerarquizan presupuestalmente la evaluación o restringen la contratación de personal calificado.

Frente a estas situaciones, las agencias encuentran reducidos medios para proseguir sus intereses. Esto sería típico en dos casos: i) si no existen mecanismos de intervención de los otros actores (por ausencia de otros principales, como por ejemplo, otros órganos del Poder Ejecutivo, del Parlamento o de las familias en el gobierno de la agencia evaluadora); y ii) cuando el ministerio es *también* el primer principal de la relación educativa, o sea, el prestador de los servicios educativos⁸.

Si la hipótesis del origen del déficit institucional es correcta, debiera ser posible observar que los países donde *no hay evaluación periódica ni difusión pública de resultados*, se caracterizan también por tener

⁶ En nuestro estudio comparado sobre México y Uruguay (Fernández and Midaglia, 2003 y 2005), el objeto fue la modificación de la relación educativa. Sin embargo, a poco de comenzar emergió con gran preeminencia en ambos países, el problema de la relación evaluativa entre las agencias y el principal. Basta recordar que en ambos países, la autoridad prohibió informar sobre al menos un proyecto de evaluación.

⁷ Las mediciones periódicas pueden reportar que los avances educativos durante una gestión no han sido espectaculares y esto conllevar a una pérdida de imagen del Ministro o de todo el Gabinete frente al electorado.

⁸ V.g. la prestación de los servicios no está descentralizada a las entidades, como en México o Argentina, o a los municipios con en el caso de Chile o Brasil.

una agencia cuyo principal es a la vez prestador del servicio (hipótesis 1), y sólo cuentan con *un único* principal en la relación evaluativa (el ministerio).

Para concluir este apartado, quiero resaltar que este enfoque no es novedoso. Ferrer (2006: 14) sostuvo en su balance de mediados de la década que: *"En cualquier caso, lo que resulta imprescindible es definir con claridad el grado de autonomía funcional que tendría la unidad evaluadora. El marco institucional puede tener fuertes implicancias para la credibilidad y legitimidad de sus acciones, para la sostenibilidad técnica-operativa y financiera del sistema, o para las posibilidades de difusión, uso e impacto de los resultados de las evaluaciones."* Precisamente, este componente de la institucionalidad de la evaluación, su gobierno, es a mi juicio, la clave que termina por calificar, mejor, explicar su fortaleza o debilidad; este artículo busca aportar evidencia empírica en esta dirección, la cual hasta el presente es inexistente.

2. EL ESTUDIO

El análisis empírico se realiza sobre la base de un relevamiento de información realizado originalmente en 2006, y actualizado al presente año (2010)⁹. El período histórico de observación estuvo comprendido entre 1980 y el año 2010; un total de treinta años. Tal ventana de tiempo tiene justificación casi exclusivamente para el registro del comienzo de las evaluaciones externas y periódicas en un país: Chile. En varios documentos, también se ha dejado constancia que México comenzó las evaluaciones en la década de 1970; sin embargo, no se ha podido registrar ninguna evaluación entre ese año y 1991. Para el resto de los países, la historia de la evaluación de aprendizajes comienza en los años noventa.

La unidad de registro de la información fue definida en la "evaluación". Se considera que un proyecto es una única evaluación cuando se realiza en un mismo ciclo escolar, durante un período determinado de meses en el año, con un mismo propósito general (débil o fuerte, en el sentido de Ravela, 2001), bajo un mismo diseño muestral, con un mismo marco conceptual, una misma teoría de la medición (clásica o TRI), aunque estuvieran dirigidos a varios grados, ciclos o niveles del sistema educativo. Bajo esta definición, un proyecto generalmente implica un operativo en el terreno, aunque puede darse el caso de que a través de una misma logística se apliquen dos proyectos, como ha sido el caso de Prova Brasil y SAEB.

Es clásica ya la distinción entre evaluaciones con alto impacto y con bajo impacto, formulada por Ravela (2001). Visto el marco teórico y los objetivos del análisis, tomé la decisión de sólo considerar evaluaciones externas que no tuvieran consecuencias directas sobre la trayectoria académicas formales de los estudiantes. En consecuencia, excluí aquellas cuya finalidad principal *fuera* la acreditación de un grado o de un nivel, tal como es el caso de la Prueba de Suficiencia Universitaria en Chile, los Exámenes de Estado en Colombia, las Pruebas de Conocimiento de Costa Rica, las pruebas PAES de El Salvador y las Pruebas Nacionales de la República Dominicana.

Con base en el marco analítico definido por dimensiones extraídas de un enfoque institucionalista combinado con la teoría agente/principal, se establecieron las siguientes dimensiones de análisis: i) definición institucional (nombre, agencia, dependencia, normativa reguladora); ii) levantamientos (años, grados, asignaturas); iii) diseño metodológico (enfoque de la prueba, tipo de reactivos, teoría de la

⁹ Agradezco la colaboración de la Lic. Inés Malán en la confección de la base de datos de 2006.

medición, equiparación, muestra); y iv) uso y difusión de resultados (destinatarios, tipos de informe, acceso a microdatos). Cada evaluación fue catalogada en estos aspectos, al menos, hasta donde fuera posible hallar información.

Cuatro fueron las fuentes principales de información, todas ellas secundarias, que fueron empleadas: i) los reportes publicados (cuando los hubo); ii) los estudios sobre los sistemas de evaluación; iii) las páginas web de PREAL, la OEI, los ministerios y las agencias de evaluación¹⁰; y iv) el Foro de Evaluación que abriera Ferrer en la página de PREAL con base en su libro *“Los sistemas de evaluación educativa en América Latina”*. La última actualización de sus fichas es para el mes de abril de 2009.

2.1. Casos seleccionados

Se construyó una base de datos para los países de la región latinoamericana¹¹, conteniendo información descriptiva sobre cada una de las evaluaciones nacionales estandarizadas realizadas. De los 19 países de la región, sobre dos no fue posible acceder a información secundaria precisa sobre las evaluaciones (Venezuela y Cuba). Otros dos, Costa Rica y República Dominicana, sólo realizaron evaluaciones con la finalidad de acreditación, hasta donde se tiene conocimiento, por lo que también fueron excluidos de los análisis. Un punto importante para el relevamiento fue establecer la vigencia de las evaluaciones. Esto se determinó genéricamente considerando si se habían realizado evaluaciones los dos últimos años, o si al menos, se habían producido informes de evaluaciones realizadas antes de 2009. Aquí resaltan tres casos: Honduras, Bolivia y Paraguay. En el primero, se identificó que la última evaluación fue en 2008 pero no hay información sobre su procesamiento. Es de recordar que el país vivió un golpe de estado en junio de 2009 y pervive aún el enfrentamiento civil desatado. Bolivia por su parte, aplicó por última vez en 2000, y Ferrer (2006) da cuenta que el sistema de evaluación conocido como SIMECAL fue cerrado en 2004, sin que haya sido re-emplazado por otro. Paraguay por su parte, realizó en 2006 la última aplicación conocida de su programa denominado SNEPE, sobre el cual se han publicado informes hasta diciembre de 2007. No he registrado actividad del SNEPE desde ese entonces.

Nicaragua fue incluida aunque presentó la siguiente ambigüedad: por un lado, no aplica evaluaciones basadas en el denominado SNE desde el año 2002; por otro, en 2008 se registró una evaluación diagnóstica hecha por un consorcio entre el Ministerio y la Universidad Nacional Autónoma aplicado al final de la Educación Media sobre el cual hay información parcial de resultados difundidos en 2009.

En síntesis, el análisis cubre el período de 30 años para Argentina, Brasil, Colombia, Chile, Ecuador, El Salvador, Guatemala, México, Nicaragua, Panamá, Perú y Uruguay.

2.2. Niveles de análisis e indicadores

El objeto mismo de la investigación requiere estudiar la evaluación educativa en términos institucionales para lo cual se requiere *inferir* si más allá de la realización de operativos o proyectos de evaluación, se han establecido haces de normas capaces de regularizar el comportamiento de los actores y redistribuir la información. En consecuencia, fue necesario establecer criterios para distinguir qué evaluación fue un emprendimiento excepcional, puntual, sin continuidad, de aquellas otras evaluaciones que formaban

¹⁰ http://www.preal.org/Grupo.asp?Id_Grupo=3&Id_Seccion=29&Id_Seccion2=139

<http://www.oei.es/quipu/evaluacion.htm>

¹¹ La región “América Latina” aquí refiere los siguientes países: Argentina, Bolivia, Brasil, Colombia, Costa Rica, Cuba, Chile, República Dominicana, Ecuador, El Salvador, Guatemala, Honduras, México, Nicaragua, Panamá, Paraguay, Perú, Uruguay y Venezuela.

parte de una política de largo plazo. El concepto de "programa" fue utilizado para marcar dos grados de institucionalización.

El "nombre" oficial de la evaluación fue el primer criterio. La pretensión de continuidad está generalmente asociada al uso recurrente de los términos "sistema", "programa" u "operativo". Es el caso, por ejemplo, del Sistema de Medición de la Calidad (SIMCE) de Chile y el Programa de Estándares Nacionales (EN) de México. En cambio, el término está ausente en los proyectos hechos sin pretensión alguna de continuidad, como fue el caso de la Evaluación en el Primer Nivel de Escolarización hecha por Uruguay en 2001 (ANEP, 2002). Sin embargo, el nombre como indicador tiene limitaciones. Un cambio de nombre en el tiempo no implica necesariamente un cambio de programa, sino más bien un cambio en la agencia de evaluación. Es el caso del término "Pruebas Nacionales" para los ciclos 2002-2003 y 2003-2004, que adoptara el Instituto Nacional de Evaluación de la Educación de México (INEE) para designar los levantamientos del Programa Estándares Nacionales antes realizados por la Dirección General de Evaluación (DGE) de la Secretaría de Educación Pública.

A su vez, la falta de nombre no indica necesariamente ausencia de programa. Algunas evaluaciones fueron tituladas simplemente "pruebas" o "evaluaciones", aunque luego generaron dos, tres y hasta más proyectos. El segundo criterio que apliqué introduce la idea *metodológica* de la replicación: sin perjuicio del nombre, habrá programa toda vez que dos o más evaluaciones sucesivas se hayan dirigido a los mismos grados y hayan utilizado una metodología similar ("tabla de especificaciones" y teorías de la medición). Uruguay cae dentro de esta categoría ya que sus evaluaciones en 6tos. Primaria de 1996, 1999 y 2002 conformaron un programa en este sentido. En Perú, las evaluaciones de 2001, 2004 y 2008 informan de una continuidad clara, auto referida en las publicaciones, pero sin apelar al término Programa. La noción utilizada fue bastante laxa: alcanzó con que dos evaluaciones tuvieran *explícitamente* esta lógica para *inferir que hubo(hay) programa* y por tanto, indicar un grado de institucionalización *incipiente* de la evaluación educativa.

Tal como lo establece la hipótesis 2A, es razonable sostener que un período más prolongado de existencia de un mismo programa indicaría mayor aceptación, conformidad, consentimiento y utilización de las reglas de juego. En consecuencia, adopté el criterio de que si hubo 3 o más operativos que siguen en la lógica de la replicación, sería un indicador de clara institucionalización de la evaluación. El criterio obedece a que en la región los mandatos presidenciales duran 4 años (Chile desde 2006, Argentina, Guatemala, Paraguay, Colombia, Brasil, Honduras, Perú), 5 años (Uruguay) o 6 años (México). El diseño de una evaluación, una vez decidida, demanda un año adicional previo al ciclo escolar de su aplicación. Tres años de levantamientos (si fueran consecutivos) y uno primero de preparación totalizan cuatro años calendario, que lindan con esos mandatos.

Ahora bien, el uso del mismo nombre puede ocultar discontinuidades importantes en uno o varios aspectos *metodológicos en la elaboración del diseño de la política* y por tanto afectar también la formación institucional (hipótesis 3). En los hechos, tanto el SIMCE en Chile como el ONE en Argentina, la PAES de El Salvador, el SAEB de Brasil, conservan programas de largo aliento pero en ambos casos han realizado al menos un cambio en la metodología que generaron cambios importantes en las reglas de juego para los actores. Estos antecedentes demandaron un cuarto criterio, a saber, examinar la existencia de una metodología *consistente* de equiparación de pruebas a lo largo del tiempo; un tema teóricamente complejo que demanda un elevado nivel de formación pedagógico-metodológico-estadístico de los técnicos que trabajan en las agencias de evaluación. Sin embargo, este registro requeriría realizar evaluaciones de las evaluaciones, una cuestión que por el momento está fuera de las posibilidades de

este trabajo. En consecuencia, el indicador utilizado es muy rudimentario e informa que un programa es metodológicamente consistente si expresamente se declara que se han conservado, de un proyecto a otro, las mismas operacionalizaciones de los objetivos o estándares educativos ("tabla de especificaciones") y la misma teoría de la medición¹². Ambos son elementos imprescindibles para poder realizar en forma válida y confiable una equiparación de las pruebas y asegurar la comparabilidad de los resultados en el tiempo. Es de destacar que el relevamiento de este indicador ha sido difícil, debido a la falta de información metodológica pública precisa que existe en muchos de los proyectos de evaluación; sólo excepcionalmente se indica que los resultados *no* son comparables en el tiempo por fallas en la equiparación de pruebas. Tal vez, esto sea, probablemente, el resultado de un comportamiento de aversión al riesgo por parte de evaluadores que podrían perder prestigio si no logran resolver en este problema metodológico tan delicado. En consecuencia, es posible que donde se identifique un programa institucionalizado en realidad haya dos, considerando la metodología. Así, la base de datos permite distinguir cuatro categorías de evaluaciones, ordenadas de acuerdo a los objetivos del trabajo, donde la consistencia metodológica en el tiempo es indicador de su robustez científica, en tanto que la pervivencia por tres o más operativos será indicador de institucionalización.

- proyectos realizados por única vez
- programas de evaluación (2 aplicaciones con los mismos objetivos)
- programas institucionalizados (3 y más evaluaciones)
- programas institucionalizados y consistentes metodológicamente

Se comprende inmediatamente que la noción de programa que aquí se utiliza es bastante amplia. En términos ideales, buena parte de la bibliografía regional ha expresado la recomendación de que los países adopten un programa de evaluación, entendido como una institución educativa, científicamente rigurosa y burocrático establecida, cuyo objetivo es la producción periódica de información sobre los niveles y contextos de aprendizajes de los alumnos de la educación básica. En consecuencia, es factible que el número de programas históricamente desarrollados en la región haya sido sobreestimado por la metodología aquí empleada. El lector deberá tener presente esta limitación a la hora de valorar las inferencias y conclusiones plateadas.

3. PRINCIPALES HALLAZGOS

Este capítulo expone los hallazgos del relevamiento realizado en 15 países de la región. Ha sido ordenado en cuatro grandes apartados: evaluaciones nacionales, evaluaciones internacionales, distribución de la información y agencias. La motivación que guía la exposición es evaluar la razonabilidad de las tres hipótesis que se introdujeron más atrás para explicar el estado actual de la evaluación en la región; brevemente: la metodológica, la política y la institucional.

3.1. Evaluaciones y programas nacionales

El cuadro 1 presenta un primer conjunto de indicadores que caracterizan la evaluación en América Latina según el número de operativos realizados y programas erigidos. En la última fila del cuadro que contiene

¹² No es necesario que el diseño muestral se mantenga de un proyecto a otro.

resúmenes, se puede apreciar que en estos 17 países y en los últimos 30 años, se estima¹³ que se realizaron 186 evaluaciones de aprendizaje; un promedio de 12 evaluaciones por país.

Habría tres grupos de países con diferente nivel de desarrollo en este campo. México es el país más evaluaciones ha realizado (43). Junto con otros cuatro países, Chile, Brasil, El Salvador y Argentina, concentran 110 evaluaciones que representa el 60%. Si se recorre la columna [1] del cuadro, se aprecia que si bien todos los países de la región han realizado evaluaciones, Nicaragua tiene la menor experiencia acumulada (3 evaluaciones), Honduras (5), Panamá (5) y Ecuador (7). Estos países a su vez, implementaron sólo en un sentido limitado las reformas de segunda generación.

Tal como se aprecia, la distribución por país parece acompañar la distribución esperada según la hipótesis 1. Chile municipalizó la educación pública y privatizó un tercio de sus escuelas, generando un sistema de vouchers en 1982. México y Argentina descentralizaron a las entidades federativas la educación básica entre 1992 y 1993. Brasil municipalizó la educación en 1988. Tres países no concuerdan con la teoría. El Salvador, Perú y Uruguay tienen una importante cantidad de evaluaciones pero que durante los años 90, no generaron una reforma estructural de tipo descentralizadora¹⁴. A su vez, Bolivia (4) ensayó programas de segunda generación importantes, sobre todo durante los gobiernos de Sánchez de Lozada. Sin embargo, la evaluación educativa no tuvo gran desarrollo en ese país en el que prosperó el gobierno unitario nacional.

3.1.1. Institucionalización incipiente y programas institucionalizados

Se aprecia que aquellas 186 evaluaciones se han realizado en su gran mayoría (179) conforman proyectos replicados al menos en una ocasión, esto es como programas, configurándose un grado de institucionalización incipiente de interés. En el otro extremo, la columna [2] del cuadro informa el número de evaluaciones *no replicadas* que hubo en la región y destacan dos países, Nicaragua y Guatemala. Este dato correlaciona bien con la mera contabilidad. Pocas evaluaciones abonan una precaria institucionalización. Sin embargo, Uruguay constituye un caso desviado, ya que de sus 17 evaluaciones, 5 fueron no replicadas: (i) la evaluación en 3ero. Primaria de 1998; (ii) la evaluación en final del Ciclo Básico de Educación Media¹⁵; (iii) la evaluación en el primer ciclo de escolarización que comprendió nivel 5 años de Inicial y 1ero y 2do Primaria en 2001; (iv) la evaluación en el último año de la Educación Media Superior en 2003; y (v) la evaluación on-line en Primaria en 2009.

Habrían existido o existen actualmente, 39 programas nacionales de evaluación externa de aprendizajes en la educación básica. Esto parecería confirmar la hipótesis general de formación de las instituciones por la estabilidad (pervivencia) de las reglas de juego. Sin embargo, una más refinada discusión de la hipótesis 2a, requiere distinguir entre programas incipientes e institucionalizados. De los 39, 8 programas sólo se aplicaron 2 veces y pertenecen a Ecuador, El Salvador, Nicaragua, Panamá, Perú y Uruguay. Estos programas registrados responden a tres períodos y situaciones distintas. Por un lado, están APRENDE de Nicaragua y la Evaluación Nacional en 6tos Primaria de Uruguay: se he hicieron en los noventa y se descontinuaron luego. Por otro lado, están el SINECA de Panamá, CRECER de Perú y SINEA (2001 y 2003) de

¹³ No puede eliminarse la probabilidad de error en las medidas hechas en cada país. Debe recordarse que este estudio realizó su relevamiento sobre la base de fuentes secundarias públicas.

¹⁴ Es discutible la ubicación de El Salvador en este esquema. El sistema, en su concepción normativa, deja en manos del Ministerio de Educación el papel de administrador, supervisor y financiador del sistema educativo formal público. Sin embargo, el sistema EDUCO generó espacios amplios de cooperación público-privado; otros instrumentos también estuvieron en el camino de la flexibilización del control centralista del Ministerio.

¹⁵ En términos internacionales, noveno grado.

El Salvador. Son casos de programas realizados en los noventa y *sustituídos* por otros en los dos mil.

CUADRO 1. CARACTERIZACIÓN DE LOS PAÍSES SEGÚN LAS EVALUACIONES DE APRENDIZAJE EN EDUCACIÓN PRIMARIA Y MEDIA REALIZADAS ENTRE 1980 Y 2010

Países	[1] N° evaluaciones nacionales	[2] N° proyectos con única aplicación	[3] = [4] + [5] N° total de programas	[4] N° Programas con sólo 2 levantamientos	[5] N° programas institucionalizados	[6] N° programas institucionalizados y consistentes	[7] Programa(s) vigente(s) (2009-2010)	[8] Grados Evaluados por Programa(s) vigente(s)
Argentina	12	0	2	0	2	2	One	3°, 6°, 9°
Bolivia	4	0	1	0	1	0		
Brasil	14	0	4	0	4	3	Saeb, Prova, Enem	4° y 8°, 3°M
Chile	26	0	3	0	3	3	Simce	4°, 8° y 2°Medio
Colombia	8	0	1	0	1	1	Saber	5°, 9° y 11°
Ecuador	7	0	2	1	1	1	Ser (**)	4°, 7°, 10° y 3°M
El Salvador	15	0	4	1	3	3	"paesita"	1°, 3°, 6° y 9°
Guatemala	10	1	2	0	2	0	Graduandos	1°, 3°, 6° y 3°M
Honduras	5	0	1	0	1	1		
México	43	0	7	0	7	6	Excale Enlace	3°, 6° y 9° 4°, 5°, 6°, 7°, 8° y 9°
Nicaragua	3	1	1	1	0	0		
Panamá	5	0	2	1	1	sd	Sineca (**)	3°, 6°, 9° y 12°
Paraguay	8	0	1	0	1	sd		
Perú	9	0	3	1	2	1	Evaluación Nacional	Pre-escolar, 2°, 4°, 6° y 2° Medio
Uruguay	17	6	5	3	2	1	Evaluación Nacional (**)	6to Primaria (**)
Totales	186	8	39	8	31	22	14	

Fuente: elaboración propia

En Ecuador se observa la puesta en marcha de una nueva política de estado. El Sistema Nacional de Evaluación y Rendición Social de Cuentas (SER) emerge de la Consulta Popular sobre el Plan Decenal de Desarrollo Educativo realizada en 2006. En su sexta línea establecía el "mejoramiento de la calidad y equidad e implementación de un sistema nacional de evaluación y rendición social de cuentas del sistema educativo". El SER incluirá cuatro componentes: la evaluación del desempeño de los estudiantes, del desempeño de los docentes, de la gestión institucional y la evaluación de la aplicación del currículo (Ministerio de Educación, 2010). La primera evaluación se concretó en 2008 y la segunda en 2009, aunque al momento sólo hay publicados resultados para 2008.

Panamá implementó durante el gobierno de Torrijos el programa de Pruebas de Logro SINECA con dos aplicaciones, 2005 y 2008, produciendo informes generales y estudio muy preliminar de factores asociados. El nuevo gobierno que asumió el año pasado no ha informado si la próxima aplicación se realizará en 2011.

La columna [5] del cuadro 1 presenta datos para discutir la hipótesis de institucionalización por estabilidad del programa (3 y más aplicaciones). Este indicador muestra ya un primer grupo de países donde la evaluación no se ha institucionalizado o lo estuvo pero en los años noventa: Nicaragua y Bolivia respectivamente.

Dos comentarios específicos merece aquí Uruguay que muestra una débil institucionalización en estos indicadores. Ha tenido dos programas en 6tos años de Primaria debido un cambio sustantivo en el diseño del programa ocasionado principal aunque no exclusivamente por la adopción de la TRI en 2004. Las dos aplicaciones más recientes, 2005y 2009, reiniciaron la serie. También es de notar que entre los programas *institucionalizados*, dos tuvieron características limitadas: uno ya referido (6tos. Primaria 1996, 1999 y 2002); y el otro asociado a la implementación de la Reforma del Ciclo Básico (1996, 1997 y 1998). Sólo hay un programa institucionalizado *multigrado*: el que aplica la Educación Media Técnica a sus estudiantes.

3.1.2. Utilidad, actualidad y consistencia

En la columna [6] del cuadro se incluye el número de programas nacionales que a la vez califican como institucionalizados y consistentes, que serían los más robustos de todos. El número total es de 24. Bolivia, Guatemala, Panamá y Paraguay no habrían tenido este tipo de programas (al menos con la información parcial disponible); a esto se le añade Panamá, dado que no se encontró información sobre este aspecto. Es notable observar aquí que Uruguay si bien tuvo un programa con estas características, ya no lo tiene.

La vigencia de los programas aporta otro matiz al cuadro. Se aprecia en la columna [7] del cuadro que sólo 14 programas (de los 39) se encuentran vigentes en el período 2009-2010, que a su vez corresponden a 11 países. Uruguay, junto con Bolivia, Ecuador, Honduras, Nicaragua, Panamá y Paraguay, carecen *al día de hoy* de un programa de evaluación externa de aprendizajes que haya alcanzado el primer grado de institucionalización (3 y más aplicaciones).

Finalmente, la columna [8] de este cuadro 1 informa la extensión de estos programas vigentes en términos de grados y niveles evaluados. Rápidamente se aprecia otro rasgo distintivo de Uruguay: su programa (no institucionalizado) *sólo evalúa en 6tos. Primaria*. En todos los demás casos, el programa es multigrado y multinivel (Primaria y Media).

3.2. Evaluaciones internacionales

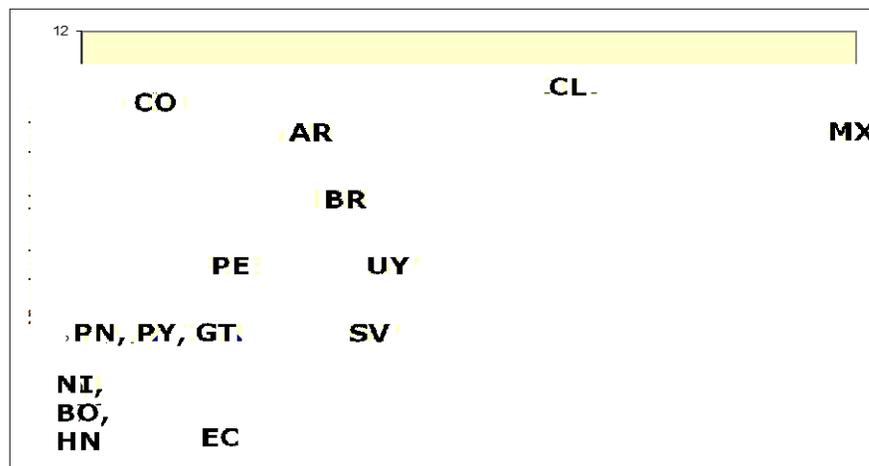
En el período observado se han realizado 13 proyectos de evaluación internacionales en los que participaron países de la región con un total de 65 operativos locales. Estos han sido llevados adelante por tres agencias: la IEA, la OECD y UNESCO. La más alta participación regional ha sido lograda por la UNESCO tanto en 1997 como en 2006, seguida por PISA ciclo 2009. (Véase en los anexos cuadro 1).

En términos de países, cinco destacan por su extensa participación: Argentina (5), Brasil (6), Colombia (9), Chile (10) y México (8). La lista prácticamente coincide con los países que concentran el mayor número de evaluaciones nacionales indicada más arriba, que tienen programas institucionalizados. La excepción aquí es El Salvador: ha participado en sólo en tres operativos (PERCE 1997, TIMSS 2007, SERCE 2006).

Esto se visualiza en el gráfico 1. Es interesante notar que, al correlacionar este indicador con el número de evaluaciones nacionales, también se descubre que también estos países presentan debilidades históricas en el plano interno (con la excepción de Ecuador), sea porque nunca tuvieron programas (Nicaragua), no los tienen vigentes (Bolivia, Honduras, Paraguay) o son de muy reciente surgimiento (Ecuador). Aquí claramente hay un cluster con cuatro países sin *programas institucionalizados vigentes* de evaluación (Bolivia, Ecuador, Honduras, Nicaragua y Paraguay), y dos países con programas institucionalizados pero con débil exposición al exterior (Panamá y Guatemala).

En el otro extremo, el gráfico 1 muestra un clúster de países larga y ampliamente evaluados, tanto en el plano nacional como en el internacional, con programas institucionalizados, consistentes y vigentes, que son Argentina, Brasil, Chile y México. Habría un cuarto grupo de países que parecerían tener un desarrollo desigual en alguna de las dos dimensiones: importante exposición internacional (Colombia) o un alto número de evaluaciones nacionales (El Salvador, Perú y Uruguay) con una participación internacional moderada y reciente (realizada en los años dos mil).

GRÁFICO1.



3.3. Población y muestra

Las constataciones anteriores permiten realizar algunas inferencias preliminares. La evaluación educativa se habría extendido en todos los países de América Latina, y habría alcanzado un primer grado de institucionalización excepto en dos países (Bolivia y Nicaragua). Esto no necesariamente conlleva tener programas robusto y menos aún, vigentes. Esta asociación parecería amparar sólo parcialmente la teoría de que la pervivencia en el tiempo es una condición muy importante para la institucionalización de los programas de la evaluación. Este apartado ensayará analizar la hipótesis 2 A, cuestión que permitirá una primera consideración de la teoría de los efectos redistributivos.

A continuación presentaré un primer aunque rudimentario indicador sobre si los programas modificaron la distribución de la información que disponen los actores externos al sistema educativo no calificados más allá de su ciudadanía. La pregunta es muy sencilla ¿faltó en alguna oportunidad un reporte público de resultados? Con base en las fuentes consultadas, la estimación preliminar se incluye en el cuadro 2.

Al menos en 6 países de la región se han registrado ausencia de información pública sobre los resultados

de las evaluaciones. Hay dos tipos de situaciones que son pertinentes destacar. Por un lado, los casos de El Salvador, México, Panamá y Paraguay, donde la ausencia de difusión alcanza a programas completos o a una parte importante de los levantamientos realizados. Las pruebas SABE de El Salvador no tuvieron reportes públicos; sólo a usuarios internos del sistema (escuelas y autoridades). El programa de evaluación aplicado en Panamá durante fines de los noventa Paraguay habría realizado evaluaciones entre 1996 y 2006, pero sólo se conocen públicamente resultados para el primer y el último levantamiento. México constituye el caso más extremo de todos, donde la política de la Secretaría de Educación, durante 6 años, fue no difundir evaluaciones de ninguno de los 4 programas nacionales que tenía implementados ni tampoco de la evaluación internacional de TIMSS 1994-1995.

CUADRO 2. CLASIFICACIÓN DE LOS PAÍSES SEGÚN SE REGISTREN CENSURAS A LA PUBLICIDAD DE RESULTADOS DE SUS EVALUACIONES

Países que restringieron más de una evaluación:	Países que no registran censuras:
<ul style="list-style-type: none"> ▪ El Salvador (Hasta 2001) ▪ México (hasta 2000) ▪ Panamá (hasta 2001) ▪ Paraguay ▪ Uruguay (2003 y 2007-2009) 	<ul style="list-style-type: none"> ▪ Bolivia (período SIMECAL, hasta 2000) ▪ Brasil ▪ Colombia ▪ Chile ▪ Perú ▪ Ecuador
Países que restringieron al menos una publicación o la demoraron:	Países sin información concluyente:
<ul style="list-style-type: none"> ▪ Argentina (ONE2003) 	<ul style="list-style-type: none"> ▪ Guatemala

En el caso de Argentina y de Uruguay, la falta de información se refiere a coyunturas muy específicas. La crisis política de los años 2001 a 2003 (renuncia del Presidente de la Rúa), llevó al cierre del Instituto de Desarrollo de la Calidad de la Educación (IDECE), generó una gran inestabilidad en la política educativa que llegó incluso a suprimir el Ministerio de Educación de la Nación durante la breve presidencia de Rodríguez Saa. En ese período, Argentina no publicó informe nacional para PISA 2000 ni para PIRLS 2001. En mayo de 2003 asume como Presidente Kirchner y como Ministro de Educación, Filmus, quienes continuaron la implementación del ONE. El informe de resultados vio demorada su publicación y tuvo un carácter muy general, no reportó por provincias ni hizo análisis más sofisticados, tal como se había venido haciendo con anterioridad. En Uruguay, la evaluación realizada en el último grado de Bachilleratos en octubre de 2003 tenía listo el informe para 2004, pero las autoridades de ese momento consideraron inoportuno su publicación¹⁶. El informe fue publicado una vez que las nuevas autoridades asumieron en 2005, aunque tuvo reducida difusión en la opinión pública y sus resultados no fueron utilizados en ningún marco (académico o político).

Cabe anotar aquí también que algunos programas pusieron a disposición de usuarios externos, las bases de datos de las evaluaciones. Es el caso del ONE, que en 2000 "colgó" las bases de 1997 a 1999, y posteriormente también la del año 2000. De hecho, Argentina fue el primer país, junto con Perú, en darle publicidad on-line a los microdatos, política que ha sido seguida por México y Brasil. Tal nivel de publicidad es más bien excepcional. Chile y Uruguay tienen políticas reguladas de acceso a los microdatos

¹⁶ Extrañamente, también censuraron otra publicación: una síntesis detallada de los logros alcanzados luego de 12 años de reformas en la educación entre 1992 y 2004.

para académicos¹⁷. Los restantes países, hasta donde se pudo llegar con este relevamiento, no existen protocolos públicos para el acceso a los microdatos por parte de académicos.

Entre los países centroamericanos se encuentra al menos un informe de difusión en El Salvador, Honduras y Guatemala, sea en formato papel o en medios electrónicos. En el primer caso, la publicidad sobrevino con las pruebas que el SINEA comenzara a aplicar en 2001 y está innegablemente asociada al uso de acreditación que tiene el PAES. Honduras ha realizado informes descriptivos, estudios de factores asociados y de tendencias en el tiempo para sus sucesivos levantamientos, y resultó hasta 2008 el país con mayor nivel de estudios de la sub-región. La prueba Graduandos de Guatemala tiene difusión en web desde 2007 y la evaluación en Primaria realizada en 2004 alcanzó amplia difusión, principalmente “empujada” por la USAID.

3.4. Agencias

La realización de la primera evaluación de aprendizajes requiere el reclutamiento de dos equipos: uno de especialistas en curriculum que definan y validen los contenidos incluidos en cada ítem de las pruebas, y otro equipo con conocimientos sólidos en muestreo, medición y análisis estadístico. También es necesario conformar una unidad de logística que se ocupe de los traslados y de la aplicación externa en cada escuela del país, siguiendo los estrictos protocolos que aseguran la validez y confiabilidad científica de las mediciones. Los estudios de caso hechos sobre este tema coinciden en señalar que son competencias profesionales escasas en casi todos los países de la región, con lo que se puede afirmar razonablemente que son muy altos los “costos hundidos” que conlleva este primer operativo. Más aún, el diseño organizacional de una agencia como unidad especializada (dentro o fuera de la administración pública), que realice sucesivas evaluaciones.

Por consiguiente, era de esperar que el relevamiento sobre el número de agencias de evaluación que han existido en gruesamente desde inicios de los noventa fuera muy bajo, próximo a 1 y que la región no hubiera constituido más de 15 o 16 agencias. Sin embargo, el panorama resulta bastante distinto: en las dos décadas pasadas se erigieron 30 agencias con un promedio de 2 por país. Esto es, una agencia por década. La moda de la distribución también es 2, habiendo tres países (Argentina, Ecuador y Guatemala) con tres agencias y 5 países que tuvieron sólo una hasta donde se ha podido relevar (Bolivia, Brasil, El Salvador y Paraguay). Uruguay se destaca en la región por ser el país que ha instituido en estos años mayor cantidad de agencias de evaluación: 5 en total.

La agencia más antigua de la región es el Instituto Nacional de Estudios y Pesquisas Anísio Teixeira (INEP) de Brasil fundado en 1937 aunque su diseño institucional actual data de 1995. Le sigue la Dirección General de Evaluación (DGE) de México creada en 1970, y el Instituto Colombiano de Fomento a la Educación Superior (ICFES) creado en los años ochenta. Las agencias más nuevas que se han creado están en Perú (Ley General de Educación N° 28044), el Instituto Peruano de Evaluación, Acreditación y Certificación de la Calidad de la Educación Básica (IPEBA) y el Instituto Nacional de Evaluación Educativa (INEE) de Uruguay (Ley de Educación n°18437), aunque este último aún está fuera de funcionamiento por razones políticas y no ha realizado ninguna evaluación (Fernández, 2009).

Por lo general, en su origen las agencias fueron gestadas en dos grandes modalidades: i) como componentes de evaluación en proyectos de mejoramiento de la educación básica financiados por

¹⁷ La excepción ha sido con los microdatos de PISA, ciclos 2003 y 2006, donde la DIEE liberó on-line la base incluyendo las variables “nacionales”.

organismos internacionales (el Banco Mundial, el Banco Interamericano, la Agencia para el Desarrollo de Estados Unidos), o ii) como consorcios inter-institucionales nacionales en los que participa al menos una universidad local. Dependiendo de los países, de esta primera agencia se pasa a una agencia incorporada a los ministerios de educación como unidades especializadas, en general, ubicadas dentro de las subsecretarías o direcciones de planificación. Este es el formato con mayor predominancia actualmente. La excepción a este formato se encuentra en Colombia con el ICES, en Brasil con el INEP y en México con el INEE, todos organismos con personería jurídica propia bajo la forma de entidades paraestatales o empresas públicas¹⁸.

En varios casos, las agencias cambian de posición en la jerarquía ministerial, aunque sin una transformación sustancial de sus funciones y dependientes. Por ejemplo, Argentina alojó a los ONE desde 1993 en la Dirección Nacional de Evaluación (DNE) dentro de la Subsecretaría de Promoción y Evaluación de la Calidad que a su vez pertenecía a la Secretaría de Programación del Ministerio; en algunos períodos, fue más bien la Subsecretaría la agencia. Durante la gestión de Lach en el Ministerio se impulsó la creación de una agencia autónoma del Ministerio que incorporara más sistemáticamente y con mayor peso la opinión de las provincias expresadas en el Consejo Federal de Cultura y Educación. El proyecto se concretó a comienzos del 2001 con la creación del Instituto para el Desarrollo de la Evaluación (IDECE). Sin embargo, la crisis política de diciembre de 2001 también arrastró al IDECE, que fue cerrado en 2002, volviendo las tareas de evaluación a ser desarrolladas por una Dirección dentro del Ministerio (actual DINIECE) (Nadorowski, Nores and Andrada, 2002: 20-24).

En la región, sólo hay un país, México, que tiene dos agencias de evaluación activas (INEE, DGE), realizando simultáneamente evaluaciones, ambos programas institucionalizados y públicos (EXCALE y ENLACE, respectivamente). Perú, a partir de 2008 creó el IPEBA pero hasta el presente no ha realizado evaluaciones y según información parcial, no las hará, manteniéndose la UMC como agencia evaluadora¹⁹.

En este contexto, destaca Uruguay donde existen actualmente tres agencias de evaluación: la Dirección de Investigación, Evaluación y Estadísticas (DIEE), la Dirección de Planeamiento de UTU, ambos dependientes de la Administración Nacional de Educación Pública (ANEP) y el creado por ley pero aún no instalado INEE.

3.5. Alternancias y evaluaciones

Cuatro países carecen de programas en la actualidad: Bolivia, Honduras, Nicaragua y Paraguay. Nicaragua nunca tuvo un programa. De hecho, ninguno de los cuatro países parecería tener hoy en día una agencia activa para la evaluación. En el otro extremo, México y Uruguay tienen dos agencias activas, cada una con evaluaciones propias, pero en una situación muy diferente como se ha revisado: el primero con 2 programas institucionalizados y el segundo con programas sin institucionalizar. Estos países, junto con Argentina y Venezuela presentan expedientes de cambios en los partidos de gobierno.

Los cuatro primeros países han tenido alternancia de partidos en el poder junto con cambios drásticos de orientación en las políticas generales de los nuevos gobiernos (Morales, Zelaya, Ortega y Lugo), siendo tal vez Honduras el caso más extremo debido a la altísima inestabilidad política en que se procesó este giro, con el consiguiente golpe de Estado, estado de violencia superviviente aún y nuevo giro político con la

¹⁸ El INEP fue erigido en 1937 dentro del Ministerio. En 1972 fue rediseñado alcanzado el estatuto de organismo autónomo, que mantuvo hasta 1985.

¹⁹ Comunicación personal de Guillermo Ferrer, julio 31 de 2010.

asunción de Lobo. Los cambios además, se han verificado en fechas próximas a la detención de los programas de evaluación. La misma asociación podría aplicársele a Venezuela, aunque Chávez asumió en 1999 y la última aplicación del SINEA de que se tiene noticia habría sido probablemente en 2003 (Ferrer, 2003).

Otros cambios de partido han tenido efectos de reafirmación, aunque no sin conflictos. La alternancia tuvo efectos drásticos en Argentina con la caída de De la Rúa y el interinato de Rodríguez Saá, pero a diferencia de los anteriores, Kirchner estableció la DINECE y el primer ONE de la nueva era se realizó ya en 2003. Algo similar debiera decirse respecto al Perú con la caída de Fujimori y el gobierno de Toledo. Chile por su parte, cambió no solo de partido o de orientación, sino también de régimen político en 1990, conservando y profundizando el entonces incipiente programa SIMCE, iniciado en 1988.

El caso de México es más especial aún. El año 2000 marcó la alternancia en el gobierno entre el Partido Revolucionario Institucional, en el poder por casi 70 años, y el Partido de Acción Nacional. Si bien en términos ideológicos fuertes, no hubo un cambio profundo en las orientaciones de ambos gobiernos, sí se produjo una modificación fuerte referida a las condiciones institucionales de la evaluación: en 2002 el Presidente Fox creó por Decreto el INEE como organismo autónomo y le asignó la competencia de evaluar la educación básica (inicial, primaria y secundaria, general y técnica).

Sin embargo, otros tantos países han experimentado alternancias de partido con en el gobierno con orientaciones contrapuestas: Brasil, Ecuador, El Salvador, Guatemala y Uruguay. En todos ellos, asumió el gobierno un partido de izquierda sin que esto hubiera significado una suspensión de los programas de evaluación existentes. En El Salvador, el gobierno ha pasado desde la Alianza (derecha) al Frente Farabundo Martí, la ex-guerrilla, sin que esto suspendiera la periodicidad de los programas PAES; es más, continuó con la instrumentación de la PAESITA. En Uruguay, la izquierda en el poder dio una señal de afirmación parlamentaria sobre la necesidad de contar con evaluaciones externas en mejores condiciones institucionales, creando el INEE como persona de derecho público no estatal, autónoma de la autoridad educativa.

Es claro que la alternancia de partidos de gobierno con orientaciones disímiles tiene efecto sobre la política de evaluación, y más en particular, sobre la supervivencia de los programas o sobre condición institucional de las agencias. Sin embargo, algunos ejemplos muestran que salvo extremos ideológicos pronunciados (tal vez Bolivia), estos giros restringen pero no anularían totalmente la evaluación, en particular en aquellos casos en que ha alcanzado un importante periodicidad y volumen: Argentina, Chile, Brasil, El Salvador, México, Perú y Uruguay. Incluso se podría augurar que una vez superado el enfrentamiento civil generado por el golpe de estado en Honduras, se retomaran las actividades de evaluación por parte de la UMCE. En mi opinión, la alternancia tiene un efecto menor sobre la institucionalización que la censura a la publicidad de la información analizada más arriba.

3.6. Principales y agencias

El análisis comparativo de la evaluación externa de aprendizajes en la región muestra diferencias significativas en el diseño institucional de las agencias, en especial, relativas al gobierno y a la dependencia funcional. En todos los países, las agencias de evaluación están vinculadas al Poder Ejecutivo, por lo general, a través del Ministerio, Secretaría o Administración de la Educación. Sin embargo, y aquí está la clave a mi entender, los países se diferencian según si la evaluación es un asunto *exclusivamente* librado a aquellas autoridades, y si esas autoridades *además*, son a su vez, prestadoras de los servicios educativos (hipótesis 4).

Con base en la teoría agente/principal, se ha elaborado el siguiente cuadro que sintetiza la información sobre dos dimensiones. La primera, ubicada en las columnas del cuadro, distingue las agencias que tienen un único principal (en todos los casos el Ministerio, Secretaría o Administración de Educación), de aquellas que tienen más de un principal, generalmente observado en la integración de la junta directiva de la agencia con miembros externos al sistema educativo. La segunda dimensión refiere a si en la estructura del sistema de educación básica se ha separado la regulación de la prestación del servicio (por ejemplo, por la vía de la descentralización o la privatización). Si ambas funciones están concentradas directamente en el Ministerio de Educación, diré que el principal (o primer principal) es prestador de servicios; en caso contrario diré que es regulador. Dentro del cuadro y para cada agencia clasificada, se añaden entre paréntesis el número de evaluaciones que ha realizado la agencia, el número de programas institucionalizados que ha tenido o tiene, y la palabra "SI" indica hechos de censura y no "NO" indica ausencia de censuras.

El resultado es bien interesante. La celda superior izquierda clasificó a las agencias cuyo principal es único y además, prestador del servicio. Ahí están Bolivia, Ecuador, El Salvador, Guatemala, Nicaragua, Panamá, Paraguay y Uruguay, totalizando 11 agencias y 68 evaluaciones. Sin embargo, estos ocho países, siete carecen de un programa institucionalizado de evaluación en marcha (la excepción es El Salvador) y cinco están actualmente vigentes. Lo más notorio es que se identifican episodios de censura en la actuación en siete agencias.

La celda inferior izquierda del cuadro informa de las agencias en las que el único principal, el Ministerio, no es el prestador principal del servicio: Chile, México, Argentina (con dos agencias) y Brasil. Es interesante notar que estos son los países ya identificados con más alto desarrollo de la evaluación educativa. De acuerdo a la hipótesis 1, la evaluación educativa debiera haber alcanzado una alta institucionalización por razones fundamentalmente sistémicas. La evidencia apoya esta idea aunque sólo parcialmente, vista la hipótesis 2B. En la mitad de las agencias relevadas, existieron episodios de censura. Destaca el historial de la DGE de México, aplicando censuras durante toda la década del noventa, a pesar de que el país implementó la federalización. La fecha de la última censura en Argentina corresponde al operativo 2003 de la DINIECE y en realidad como se dijo antes, sólo habría afectado cierto tipo de análisis.

En ambas celdas de la segunda columna, se constata que no hay casos de censura de la información, si bien es cierto que ahí sólo cuentan con tres casos en total: el INEE de México, el ICES de Colombia y la UMCE de Honduras. Dos de estos países tienen un alto desarrollo de la evaluación educativa, en tanto Honduras se encuentra en la situación opuesta. Aquí no hay episodios de censura. Esto es, con independencia de si el primer principal es prestador de servicios, parecería que la clave está en la multiplicidad de principales.

Las tres agencias tienen marcadas diferencias en la integración de su gobierno. El INEE de México está regido por una Junta Directiva de 15 miembros, titulares de organismos gubernamentales o de organizaciones civiles cuyas funciones se relacionan con los objetivos del Instituto: la Secretaría de Educación Pública (que lo preside), el Consejo Nacional de Ciencia y Tecnología, la Secretaría de Hacienda y Crédito Público, el Centro de Investigación y Docencia Económica, el Instituto Mexicano del Petróleo, la Fundación para la Cultura del Maestro, el Sindicato Nacional de Trabajadores de la Educación, el Centro de Investigación y Estudios Avanzados (Politécnico Nacional), la Universidad Pedagógica Nacional, Observatorio Ciudadano de la Educación, Consejo Rector de Transparencia Mexicana, la Unión Nacional de Padres de Familia, la Federación Nacional de Asociaciones de Padres de

Familia y la Comisión de Educación del Sector Empresarial. Esta integración corporativa es típica de la vida política en México: obsérvese que no hay representantes de los partidos políticos ni delegados del Congreso. Sin perjuicio de ello, la integración le otorga legitimidad decisoria sobre evaluación educativa a una gran pluralidad de actores sociales y burocracias públicas. En este otro ámbito, tampoco existe demasiados espacios para las acciones unilaterales y en cambio las constricciones del consenso son también evidentes. Por más que la fijación de una política nacional de evaluación sea competencia compartida entre la SEP y el INEE, esta se debe ventilar, discutir y aprobar en este órgano colegiado con otros actores ajenos a la burocracia escolar que están interesados en particular en reducir las asimetrías de información. Podríamos decir que el "principal" está claramente identificado.

CUADRO 3. CLASIFICACIÓN DE LAS AGENCIAS DE EVALUACIÓN SEGÚN PROPIEDADES DE SU PRINCIPAL

	Principal único	Varios principales
Principal prestador	UMRE, UY (3,1, NO) DIEE, UY (6, 1, SI) MESyFOD (3,1, NO) PE-UTU, UY (4,1, SI) DIR.EVL., PA (6,1, SI) SUBS.PLAN., EC (6,1, SI) DIGEDUCA, GT (10,2, ?) SNEPE, PY, (8,1, SI) SIMECAL, BO (4,1, NO) SNED, NI (3, 0, SI) DEM, SAV (15, 3, SI)	UMCE, HN (5,1,NO)
Principal regulador	UEC, CL (26,3, NO) DGE, MX (35,5,SI) UMC, PE (9, 3, NO) DNE, AR (6,1,NO) IDECE, AR (1,1, NO) DINIECE, AR (3,1,SI) INEP, BR (14,4,NO)	ICFES, CO (8,1, NO) INEE, MX (6,1,NO)
<p>Nota: Se incluye entre paréntesis: el número de evaluaciones nacionales hechas, el número de programas institucionalizados y si hubo (SI) o no casos de censura (NO) de los informes Fuente: elaboración propia con base en relevamiento realizado en 2006 y 2010.</p>		

El ICFES por su parte está dirigido por tres miembros, nombrados dos por el Presidente de la República y un tercero por el Ministro de Educación. Sin dudas, esta composición le otorga una menor representatividad a la diversidad de intereses existentes, pero cumple con darle mayor peso a los actores externos al sistema educativo, que por definición no tienen intereses contradictorios en tanto prestadores del servicio.

La UMCE por su parte nació y mantuvo hasta su última época, una naturaleza híbrida: surgida como parte de un convenio entre la Secretaría de Educación Pública y la Universidad Pedagógica Nacional Francisco Morazán, conservó un estatuto autónomo propio de la vida universitaria, más que de la vida política; dicho esto en el contexto muy relativo de la vida política hondureña.

Vale la pena indicar que el establecido INEE de Uruguay clasificaría en esta celda, dado que si bien su primer principal es prestador del servicio, su comisión directiva estaría integrada por un miembro designado por el Ministerio de Educación (que lo presidiría); tres miembros designados por la ANEP, dos

por la Universidad de la República y uno por el sector privado.

4. CONSIDERACIONES FINALES

La evidencia relevada en 15 países de la región muestra que la evaluación educativa es un fenómeno de una magnitud muy importante tanto por la cantidad de levantamientos (nacionales e internacionales) realizados en los últimos 30 años, como por el número de programas nacionales de evaluación puestos en marcha. Sin embargo, existen diferencias notorias. Por un lado, es visible un grupo de países con un alto desarrollo en la materia (Argentina, Brasil, Chile, México y Colombia). Todos han realizado extensas reformas de segunda generación en sus sistemas educativos. Por otro lado, al menos cuatro países presentan debilidades importantes debido a la inexistencia actual de programas (Nicaragua, Bolivia), la inactividad de la agencia en los últimos dos años (Honduras, Paraguay) o el reducido número de evaluaciones hechas, todas en los últimos años (Panamá, Guatemala). En todos, las reformas de segunda generación han sido muy parciales o han fracasado. Esto avalaría la primera hipótesis propuesta. Sin embargo, entre estos dos grupos de países, se ubican Perú, El Salvador y Uruguay, los que aún teniendo niveles semejantes en cuanto al número de evaluaciones, difieren entre sí tanto en la extensión de sus reformas como en la institucionalización de la evaluación. Destaca Uruguay por carecer de un programa institucionalizado *vigente*.

El examen de la hipótesis de la orientación de los partidos de gobierno y su alternancia han permitido mostrar que este fenómeno político tiene una incidencia notable en la supervivencia tanto de programas como de agencias. Sin embargo, así como hay casos notables que confirmarían la hipótesis, también hay otros que la disconfirman. Las excepciones, precisamente, parecen más relacionadas con la propia fortaleza de las instituciones de evaluación.

El análisis final practicado clasificó a las agencias según su dependencia funcional, distinguiendo cuatro posibilidades por la combinación de principal regulador / principal prestador y principal único / múltiples principales. Los tres indicadores de institucionalidad muestran un alineamiento bastante notorio: aquellos países cuyas agencias están subordinadas a un único principal y éste además es prestador del servicio educativo, tendrían mayores probabilidades de presentar episodios de censura a la distribución de la información. Dicho de otra forma, entre todas las instituciones que regulan la evaluación, la fortaleza o debilidad parecería ser el resultado de las condiciones en que es gobernada la agencia de evaluación. A pesar de que el modelo de análisis no está controlando simultáneamente el conjunto de variables que indican las hipótesis desplegadas, se puede avanzar refinando la hipótesis 4. Si bien es razonable atribuir la institucionalización de la evaluación educativa a las estructuras descentralizadas de los sistemas, y que las orientaciones liberales de los gobiernos fomentan la redistribución de información, parecería que son condiciones suficientes. El más alto grado de institucionalización se alcanza en los países que cuentan, además, con agencias no subordinadas exclusivamente al Ministerio. La conjunción de intereses, originarios de las relaciones evaluativas y prestatarias en el principal de las agencias de evaluación, parecería proporcionar una precaria institucionalidad a la evaluación educativa en términos de programas, tal como aquí se lo ha entendido. Este es un resultado muy importante en la discusión de las políticas sobre la materia. Si bien la bibliografía especializada ha destacado los aspectos institucionales, en general los ha puesto en un segundo plano a la hora de formular recomendaciones sobre cuáles son las evaluaciones que América Latina necesita. Sobre todo es un resultado fundamental para el Uruguay, país que a pesar de haber creado por Ley un Instituto autónomo, aún no ha sido

instrumento y persisten dudas sobre su contribución al mejoramiento de la calidad de la educación.

REFERENCIAS BIBLIOGRÁFICAS

- Benavidez, V. (2010). Las evaluaciones de logros educativos y su relación con la calidad de la educación. *Revista Iberoamericana de Educación*, 53, pp. 83-96.
- Benveniste, L. (2000). Student Assessment as a Political Construction: The Case of Uruguay. *Education Policy Analysis Archives*, 8. Consultado el 7 de agosto de 2010 en <http://epaa.asu.edu/ojs/article/viewFile/423/546>.
- Benveniste, L. (2002). The Political Structuration of Assessment: Negotiating State Power and Legitimacy. *Comparative Education Review*, 46(1), pp. 89-118.
- Bertolini, M. et al. (2001). *Educación Secundaria: la reforma impuesta. Días visiones críticas*. Montevideo: FENAPES / NORDAN.
- Chubb, J.E. and Moe, T. (1990). *Politics, Markets and America's schools*. Washington: The Brookings Institution.
- DIEE (2002 a). *Los niveles de desempeño al inicio de la educación primaria. Estudio de las competencias lingüísticas y matemáticas. Evaluación nacional en el primer nivel de la escolaridad: inicial 5, 1er. y 2do año. Primer informe. Setiembre 2002*. Montevideo: General de Planemiento y Gestión Educativo, Gerencia de Investigación y Evaluación / ANEP.
- DIEE (2002 b). *Evaluación nacional de aprendizajes en lenguaje y matemática. 6to año de Enseñanza Primaria. - 2002. Primer Informe de devolución de resultados de la muestra nacional*. Noviembre 2002. Montevideo: General de Planemiento y Gestión Educativo, Gerencia de Investigación y Evaluación / ANEP.
- DIEE (2004). *Informe Nacional PISA 2003*. Montevideo: ANEP. Elaborado por Ravela, P.
- DIEE (2007). Uruguay en PISA 2006. Primeros resultados en ciencias, matemáticas y lectura. Montevideo: ANEP.
- Di Gropello, E. (1999). Los modelos de descentralización educativa en América Latina. *Revista de la CEPAL*, 68, pp.153-170.
- Di Gropello, E. (2004). La descentralización de la educación y las relaciones de rendición de cuentas en los países latinoamericanos. *Documento de trabajo*, 30. PREAL: Santiago de Chile.
- Eisenhardt, K. (1989). Agency Theory: An Assessment and a Review. *Academy of Management Review*, 14 (1), pp. 57-74.
- Fernández, T. (2004). De las escuelas eficaces a las reformas de segunda generación. *Estudios Sociológicos*, 64, pp. 377-408.
- Fernández, T. (2005) Persistent Inequalities in Uruguayan Primary Education 1996 - 2002. In Teese, Richard; Lamb, Stephen and Duru-Bellat, Marie *International Studies in Educational Inequality, Theory and Policy*, 2, pp. 177-206. Dordrecht: Springer.
- Fernández, T. (2008). *El Proyecto Instituto de Evaluación Educativa del Uruguay: ¿un paso hacia la rendición de cuentas?*. Segundo Congreso Uruguayo de Ciencia Política. Montevideo.

- Fernández, T. (2009). La instalación del nuevo Instituto de Evaluación Educativa de Uruguay: desafíos organizacionales e institucionales. En Caetano, G. *Encrucijada 2010: la política uruguaya a prueba. Observatorio Político. Informe de coyuntura n°9*. Montevideo: Editorial Fin de Siglo / Instituto de Ciencia Política/CLACSO.
- Fernández, T. and Midaglia, C. (2005). *Quiénes y cómo usan los informes generados por los sistemas de evaluación de aprendizajes en la educación primaria. Los casos de México y Uruguay*. Informe Final de Investigación. México: El Colegio de México.
- Fernández, T. and Midaglia, C. (2005). El uso de los informes generados por los sistemas de evaluación de aprendizaje en la educación primaria. Los casos de México y de Uruguay. En Cueto, S. (ed) *Uso e impacto de la información educativa en América Latina*. Santiago de Chile: PREAL.
- Ferrer, G. (2006). *Sistemas de evaluación de aprendizajes en América Latina. Balance y desafíos*. Santiago de Chile: PREAL.
- Gauri, V. and Vawda, A. (2003). Vouchers for Basic Education in Developing Countries. A Principal-Agent Perspective. Washington, *World Bank Policy Research Working Paper 3005*.
- Hevia Rivas, R. (1992). *Política de descentralización en la educación básica y media en América Latina*. Santiago de Chile: UNESCO/REDUC.
- Hirshman, A. (1970). *Exit, voice and loyalty*. MA: University of Cambridge Press.
- Knight, Jack (1997). *Institutions and social conflict*. MA: Cambridge University Press.
- Martínez Rizo, F. (2010). Políticas de evaluación educativa y organismos a cargo de ellas. Experiencias y tendencias recientes en América Latina. Ponencia a las *Jornadas de Educación: una mirada a 10 años*. Montevideo: Universidad Católica del Uruguay.
- Murillo, J. and Román, M. (2010). Retos de la evaluación de la calidad de la educación en América Latina. *Revista Iberoamericana de Educación, 53*, pp. 83-96.
- Narodowski, M., Nores, M. and Andrada, M. (2002). *La evaluación educativa en la Argentina. Desde los operativos nacionales hasta los boletines escolares*. Buenos Aires: Prometeo 3010.
- North, D. (1995). *Instituciones, cambio institucional y desempeño económico*. México: Fondo de Cultura Económica.
- Postlethwaite, N. and Kellaghan, T. (2009). National Assessment of educational achievement. *Education Policy Series n°9*. Ginebra: UNESCO.
- Ravela, P. (2001). *¿Cómo presentan sus resultados los sistemas de evaluación de aprendizaje de América Latina?*. Documento de trabajo-Grupo de Estándares y Evaluación de Aprendizajes. Santiago de Chile: Grade/PREAL.
- Ravela, P., Picaroni, B., Fernández, T., Gonet, D., Haretche, C., Loureiro, G. and Luaces, O. (2000). *Evaluaciones nacionales de aprendizajes en Educación Primaria en el Uruguay (1995-1999)*. Montevideo: Ed. UMRE-MECAEP-ANEP.
- Ravela, P., Arregui, P., Valverde, G., Wolfe, R., Ferrer, G., Martínez, F., Aylwin, M. and Wolf, L. (2008). Las evaluaciones de aprendizaje que América Latina necesita. *Documento de Trabajo n°40*. Santiago de Chile: PREAL. En: http://www.preal.org/Biblioteca.asp?Id_Carpeta=64&Camino=63|Preal%20Publicaciones/64|PREAL%20Documentos.

- Ruiz, G. (2003). Proyectos de evaluación de la educación básica en la década 1990-2000. En Zorrilla, Margarita (ed.) *La evaluación de la educación básica en México 1990-2000. Una mirada a contraluz*. Aguascalientes: Universidad Autónoma de Aguascalientes.
- UMRE (1996a). *Evaluación nacional de aprendizajes. Fundamentos*. Montevideo: ANEP.
- UMRE (1996b). *Primer informe público de resultados. Evaluación Nacional de Aprendizajes en Lengua Materna y Matemática. 6to año de Enseñanza Primaria 1996*. Montevideo: ANEP.
- UMRE (1997a). *Segundo Informe público de resultados. Evaluación Nacional de Aprendizajes en Lengua Materna y Matemática. 6to año de Enseñanza Primaria 1996*. Montevideo: ANEP.
- UMRE (1997b). *Tercer Informe Público. Evaluación Nacional de Aprendizajes en Lengua Materna y Matemática. 6tos año de Enseñanza Primaria 1996*. Montevideo: ANEP.
- UMRE (1999). *Evaluación Nacional de Aprendizajes en sextos años de Educación Primaria. Primer Informe*. Montevideo: ANEP.
- Winkler, D. and Gershberg, A. (2000). Los efectos de la descentralización del sistema educacional sobre calidad de la educación en América Latina. *Documento de Trabajo n° 17*. Santiago de Chile: PREAL.
- Winters, M. (2006). La elección de escuelas en Estados Unidos. En Corvalán, J. and McMeekin, R.W. (ed.) *Accountability educational: posibilidades y desafíos para América Latina a partir de la experiencia internacional*. Santiago: CIDE/PREAL.
- Wolff, L. (2004). Educational Assessment in Latin America: the State of the Art. *Applied Psychology*, 53(2), pp.192-214.
- Zorrilla, M. (ed.) (2003). *La evaluación de la educación básica en México 1990-2000. Una mirada a contraluz*. Universidad Autónoma de Aguascalientes: Aguascalientes.

ANEXO 1

Participación en evaluaciones internacionales

TIMSS 1995	TIMSS 1999- R	LLECE 1997	CIVIC EDUC. (1994-2002)	PISA 2000 PISA PLUS	PIRLS 2001	TIMSS 2003	PISA 2003	LLECE 2006	PISA 2006	TIMSS 2007	PISA 2009	Iccs 2009	TOTAL POR PAIS
Argentina		Argentina		Argentina	Argentina	Argentina		Argentina	Argentina		Argentina		8
		Bolivia											1
		Brasil		Brasil			Brasil	Brasil	Brasil		Brasil		6
	Chile	Chile	Chile	Chile		Chile		Chile	Chile		Chile	Chile	10
Colombia		Colombia	Colombia		Colombia			Colombia	Colombia	Colombia	Colombia	Colombia	9
		Costa Rica						Costa Rica					2
		Cuba						Cuba					2
								Ecuador					1
								El Salvador		El Salvador			2
								Guatemala				Guatemala	2
		Honduras											1
México		México		México			México	México	México		México	México	8
		Nicaragua						Nicaragua					1
		Paraguay						Paraguay					2
		Panamá						Panamá			Panamá		2
		Perú		Perú				Perú			Perú		4
		Rep. Dom.						Rep. Dom.			Rep. Dom.	Rep. Dom.	3
							Uruguay	Uruguay	Uruguay		Uruguay		4
		Venezuela											1
3	1	13	2	5	2	2	3	16	6	2	9	5	69

Fuente: elaboración propia con base en el relevamiento comparado. En anexos se detallan las fuentes específicas para este punto.

ESTUDIO DIFERENCIAL DEL RENDIMIENTO ACADÉMICO EN LENGUA ESPAÑOLA DE ESTUDIANTES DE EDUCACIÓN SECUNDARIA DE BAJA CALIFORNIA (MÉXICO)¹

DIFFERENTIAL STUDY OF THE ACADEMIC PERFORMANCE IN SPANISH LANGUAGE OF HIGH-SCHOOL STUDENTS AT BAJA CALIFORNIA (MÉXICO)

Luis Lizasoain² Hernández y Luis Joaristi Olariaga

Revista Iberoamericana de Evaluación Educativa 2010 - Volumen 3, Número 3

<http://www.rinace.net/riee/numeros/vol3-num3/art6.pdf>

Fecha de recepción: 10 de octubre de 2010
Fecha de dictaminación: 27 octubre de 2010
Fecha de aceptación: 29 octubre de 2010

¹ Este trabajo ha sido realizado en el contexto del proyecto titulado *Estrategia integral 2009: factores asociados al aprendizaje*, financiado por el Sistema Educativo Estatal de Baja California y coordinado por la Unidad de Evaluación Educativa de la Universidad Autónoma de Baja California.

² En el caso del coautor Luis Lizasoain, esta investigación se realizó en el contexto de una estancia sabática durante el curso 2009-2010 en la Universidad Autónoma de Baja California para la que se dispuso de financiación mediante una subvención del subprograma de estancias de movilidad de profesores e investigadores españoles en centros extranjeros, modalidad A; concedida por el Ministerio Español de Educación.

1. INTRODUCCIÓN, CONTEXTO Y OBJETIVOS

Como es sabido, cada vez son más numerosas y frecuentes las evaluaciones a gran escala en las que se estudia y tratan de explicar las diferencias observadas en el desempeño académico de los estudiantes.

Con respecto al sistema educativo mexicano y centrados en el estudio de las variables contextuales se dispone de los siguientes trabajos: Backhoff, Bouzas, Hernández y García, 2007; Backhoff, Bouzas, Contreras, Hernández y García, 2007; Valenti, 2007; Backhoff, Andrade, Sánchez y Peón, 2008.

Centrados en el estudio de materias curriculares como las matemáticas está el trabajo de Ruíz 2009 o el de Sánchez (2009) sobre las pruebas Excale, y el de Barrera (2009) relativo a la Formación Cívica y Ética. Sobre la misma materia, pero centrado en el caso específico del Estado de Baja California son reseñables los trabajos de Carvallo, Caso y Contreras (2007) y de Caso, Hernández y Rodríguez (2009).

En esta misma línea y en virtud de un acuerdo con la Secretaría de Educación Pública del Estado de Baja California (México), la Unidad de Evaluación Educativa del Instituto de Investigación y Desarrollo Educativo de la Universidad Autónoma de Baja California ha desarrollado en el curso 2008-2009 una estrategia evaluativa integral dirigida a la población escolar de enseñanza secundaria del Estado de Baja California.

La finalidad de dicha estrategia es estudiar los factores asociados al aprendizaje y al rendimiento académico con objeto de elaborar y sugerir acciones y propuestas de mejora a los responsables del sistema educativo.

En esta línea y durante el citado curso, la estrategia evaluativa consideró la evaluación de las asignaturas de Matemáticas, Lengua Española y Formación Cívica y Ética. El rendimiento en estas materias es medido mediante las pruebas de ENLACE que se aplican con carácter censal en toda la República. Para un análisis de los factores asociados al aprendizaje del español y las matemáticas partiendo de las pruebas de ENLACE 2007 véase el ya citado trabajo de Valenti (2007).

Para la medición de los factores asociados al aprendizaje, se elaboraron cuestionarios *ad hoc* dirigidos a los estudiantes, docentes y directores. Estos instrumentos se aplicaron a una muestra representativa aleatoriamente seleccionada. Aunque en el apartado siguiente se va a describir con detalle el proceso de muestreo, adelantamos aquí que estos cuestionarios se aplicaron a casi 6000 estudiantes de tercero de educación secundaria pertenecientes a 65 centros escolares.

Aunque, como acabamos de afirmar, la finalidad general de esta estrategia es el diseño de propuestas y acciones de mejora, este trabajo se plantea como principal objetivo *el estudio de la incidencia de estos factores asociados al aprendizaje de los estudiantes de tercero de secundaria de Baja California en su rendimiento académico en la materia de Lengua Española*. Más concretamente, el interés se centra en realizar un *estudio diferencial* de dicha incidencia en función de las muy diversas situaciones y realidades detectadas en los centros escolares de Baja California. La hipótesis básica de trabajo es que la realidad social mexicana (y por ende la de Baja California) es tan diversa y tiene contrastes tan acentuados que las oportunidades de aprendizaje inciden de manera distinta en diferentes grupos de centros. Para contrastar esta hipótesis, en primer lugar vamos a caracterizar estos diferentes grupos de escuelas mediante la técnica de minería de datos conocida como árboles de decisión o árboles de regresión y clasificación. Posteriormente, mediante técnicas de modelización multinivel diseñaremos y validaremos un modelo explicativo del rendimiento en lengua tomando en consideración los resultados obtenidos en la fase exploratoria inicial.

La idea básica es que las condiciones, las oportunidades de aprendizaje relevantes no son las mismas para todos los estudiantes y todas las escuelas sobre todo si tomamos en consideración las acusadas diferencias en el nivel socioeconómico y cultural de las familias y la media de los centros. Dicho de otra manera: pensamos que, dentro de las variables que se pueden considerar como clásicas predictoras del rendimiento académico, hay algunas que inciden sobre el rendimiento de todos los estudiantes y centros mientras que otras sólo afectan a determinados subgrupos de la población. Si esta hipótesis se cumple, no basta con elaborar y ajustar modelos para toda la población, sino que debe realizarse un estudio diferencial centrado en estos subgrupos.

2. MÉTODO

2.1. Muestra

Como antes se ha señalado, la evaluación se realizó en una muestra representativa de estudiantes de tercero de secundaria seleccionados aleatoriamente. Es oportuno señalar que el Estado de Baja California consta de cinco municipios y que los centros escolares en los que se imparte la educación secundaria pueden ser de diferentes modalidades: secundaria general, técnica, telesecundaria y privada. Excepto en el caso de la privada, el resto reciben financiación pública.

Según datos del sistema educativo estatal, la distribución de la población escolar de tercero de secundaria es la que aparece en la tabla 1 en la que se observa la distribución por municipios y tipos o modalidades de escuelas del total de centros (427) y estudiantes (48.682) para este grado escolar en Baja California.

TABLA 1. ALUMNOS Y CENTROS ESCOLARES POR MUNICIPIO Y TIPO DE ESCUELA.

	GENERAL	PRIVADA	TECNICA	TELESECUN DARIA	Total
ENSENADA	4134 34	417 15	2140 14	663 37	7354 100
MEXICALI	10874 82	1118 25	2264 17	72 8	14328 132
PLAYAS DE ROSARITO	1024 5	141 4	458 2	87 4	1710 15
TECATE	1263 5	66 4	121 3	146 7	1596 19
TIJUANA	13936 58	2478 66	6223 22	1057 15	23694 161
Total	31231 184	4220 114	11206 58	2025 71	48682 427
Fuente: Sistema Educativo Estatal					

Tomando en cuenta estos estratos, la selección de las escuelas y estudiantes fue determinada mediante el método de muestreo por conglomerados en tres 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.

En función de ello, se estableció un tamaño de muestra de 71 escuelas que representa el 16.6% del total de secundarias del estado. Finalmente, de acuerdo con los registros del Sistema Educativo Estatal el número de alumnos inscritos en los grupos de tercer año de secundaria de las 71 escuelas secundarias seleccionadas fue de 8,375, lográndose encuestar solamente a 6,317 de ellos, lo que representa el 75.4 % de cobertura de la muestra aleatoria. Por último, al integrar las bases de datos de ENLACE con las de elaboración propia, el número final de centros fue de 65 y el de estudiantes de 6.211 que representan el 74.2% de la muestra estimada. En la tabla 2 se detalla la distribución de dichos estudiantes en cuanto a su sexo, edad, municipio, tipo de escuela y turno escolar. Hay que hacer notar que estos son los datos iniciales del estudio. La matriz de datos con la que vamos a operar es ligeramente distinta una vez depurada de algunos errores y de escuelas con muy bajo número de estudiantes (menos que 10).

TABLA 2. CARACTERÍSTICAS DE LOS ESTUDIANTES (N=6211)

Variable		n	Porcentaje
Sexo	Femenino	2,971	50.0
	Masculino	2,973	50.0
Edad	13	9	0.1
	14	1,649	27.4
	15	3,582	59.6
	16	679	11.3
	17 o más	91	1.5
Municipio	Ensenada	1,284	20.7
	Mexicali	1,340	21.3
	Rosarito	371	6.2
	Tecate	332	5.4
	Tijuana	2,884	46.3
Tipo de secundaria	Privada	630	10.1
	General	3,704	59.6
	Técnica	1,675	27.0
	Telesecundaria	202	3.3
Turno escolar	Matutino	4,748	76.4
	Vespertino	1,463	23.6

2.2. Variables e instrumentos de medida

Como anteriormente se ha señalado, de las tres materias cuyo rendimiento fue medido mediante las pruebas de ENLACE, en este estudio nos vamos a centrar en Lengua Española. En la página web de la Secretaría de Educación Pública del Gobierno Federal mexicano (<http://www.enlace.sep.gob.mx/gr/>) se puede encontrar información detallada sobre el programa ENLACE, las pruebas empleadas y los resultados obtenidos. Las pruebas de ENLACE son pruebas objetivas y estandarizadas, de aplicación censal y controlada. Su diseño se alinea al currículum de la educación básica del sistema educativo mexicano y pretenden informar en qué medida los estudiantes adquirieron los conocimientos y habilidades que establece dicho currículum.

Para la medición de los factores asociados al aprendizaje se diseñaron y aplicaron cuestionarios ad hoc dirigidos a estudiantes, profesores y directores. En este estudio vamos a emplear la información proveniente del cuestionario del alumno. El mismo es un instrumento de autoinforme conformado por 58 ítems que exploran las oportunidades para aprender que le ofrecen a un estudiante los diferentes contextos donde interactúa cotidianamente. El cuestionario cuenta con tres secciones: a) contexto familiar; b) antecedentes académicos; y c) contexto escolar.

Desde una perspectiva multinivel, la información proporcionada por este cuestionario nos permite recabar datos a dos niveles: estudiante y centro escolar. Para el segundo nivel se han calculado valores agregados para cada unidad del segundo nivel tomando en consideración los valores promedios de las variables relevantes.

Dichas variables han sido las siguientes:

Estudiante:

- Personales y sociodemográficas: sexo, edad, edad normativa, lugar nacimiento, lugar de residencia, años de residencia en la entidad.
- Nivel socioeconómico y cultural: indicadores sobre el nivel socioeconómico de las familias (NSE), sobre el consumo de bienes culturales, nivel de escolaridad de padres y madres, puestos de trabajo, nivel profesional. Se calcula un índice compuesto individual (familiar).
- Apoyo e implicación familiar en las tareas escolares: indicadores sobre el grado y manera en que las familias se implican y apoyan al estudiante en las tareas y sobre el nivel y forma en que motivan al estudiante. Se calcula un índice compuesto.
- Adaptación e integración escolar: Razones para estudiar la educación secundaria. Relación profesor-alumno y alumnos entre sí e indicadores sobre la calidad de la misma. Se calcula un índice compuesto.
- Lectura: horas dedicadas a la lectura, gusto y afición por la lectura, compromiso lector.
- Escolaridad: asistencia a preescolar, días que falta a la escuela, retrasos en la asistencia al centro. Horas dedicadas a hacer las tareas de español, número de tareas realizadas.
- Cuestiones sobre el docente de lengua española (percibidas por el estudiante): actividad, puntualidad, faltas, técnicas de motivación que emplea, estrategias y usos docentes.
- Equipamientos del centro: enumeración, estado y frecuencia de uso.

2.3. Preparación de datos y archivos

Una vez elaborados mediante análisis de componentes principales los índices compuestos correspondientes a los diferentes constructos, se procede a la fusión de los dos archivos de fuentes de datos parciales: pruebas de ENLACE y cuestionario de estudiantes. Este proceso de fusión de datos se realiza empleando los identificadores disponibles de estudiantes y centros. Pero aquí nos encontramos ya con una primera característica diferencial del sistema educativo mexicano. Nos referimos a la existencia de dos turnos: matutino y vespertino. En Baja California casi la cuarta parte de los estudiantes de secundaria acuden a clase por las tardes. Son numerosos los centros (todos de titularidad pública) que duplican turno y que reciben estudiantes tanto a la mañana como a la tarde.

Pero en el caso en que en un centro se dé esta situación, ambos turnos lo único que comparten, lo único que tienen en común, son las instalaciones físicas del centro escolar. El resto es completamente diferente: por supuesto los estudiantes, pero también los docentes y la dirección. Cuando en una escuela se da esta circunstancia, este hecho nos ha llevado a tomar la decisión de considerar el turno matutino y el vespertino como dos centros educativos distintos.

Así, en la tabla 3 se pueden apreciar las diferencias que se dan entre ambos turnos en variables clave como son el propio rendimiento en español, el nivel socioeconómico familiar y el nivel socioeconómico medio de las escuelas. Estas diferencias resultan significativas ($p < .01$) por lo que los estudiantes que acuden a estudiar por la mañana son de un nivel socioeconómico y cultural mayor que los de la tarde y obtienen mejor rendimiento en la prueba de ENLACE de lengua española (28 puntos).

Esto nos reafirma en la decisión de considerar a los diferentes turnos como centros distintos y para ello se elaboró un nuevo identificador de centro que incorpora el de la escuela y el del turno.

TABLA 3. DIFERENCIAS SEGÚN EL TURNO

Turno al que asiste		Calificación en Español	NSE familiar	NSE medio de la escuela
1 Matutino	N	4748	3979	4748
	Media	519,56	,05	,038
	Desv. típ.	99,41	1,02	,65
2 Vespertino	N	1463	1204	1463
	Media	491,41	-,17	-,18
	Desv. típ.	90,17	,90	,39
Total	N	6211	5183	6211
	Media	512,93	,00	-,01
	Desv. típ.	98,04	1,00	,61

Finalmente, una vez eliminados los centros con menos de 10 alumnos, el archivo de trabajo resultante contiene 6.185 estudiantes agrupados en 71 *centros-turnos* distintos. Son 13 las escuelas que duplican turno, 31 las que sólo imparten por la mañana y 4 las que en esta muestra sólo imparten por la tarde.

A continuación, y con objeto de disponer de variables del nivel de centro, se generan variables agregadas mediante el cálculo de los promedios de las variables del nivel 1 segmentando por el identificador de *centro-turno*. De esta forma, y aunque la fuente de información sea el estudiante a través del cuestionario, se dispone de variables relativas al centro-turno como el nivel socioeconómico medio de cada escuela, el promedio del clima escolar, el de las faltas de los profesores de español, etc. que completan las habituales como el turno o la titularidad.

2.4. Análisis de datos y resultados

2.4.1. Estudio mediante técnicas de segmentación

Con este archivo de trabajo se comienza el análisis empleando técnicas de minería de datos, en concreto técnicas de segmentación también conocidas como árboles de decisión o árboles de regresión y clasificación.

Se trata de un conjunto de técnicas que permiten definir y validar modelos de forma que se pueda determinar qué variables (predictoras) inciden o explican los cambios de una variable dependiente. Son técnicas estadísticas de la familia de la regresión o el análisis discriminante, pero tienen la ventaja de que tanto la variable criterio como las predictoras pueden ser de cualquier tipo (cuantitativas o cualitativas).

El principio básico de operación consiste en dividir progresivamente un conjunto en clases disjuntas. Para ello, una vez definida la variable dependiente, se efectúa una partición del grupo original en 2 ó más subgrupos atendiendo a los valores de la variable predictora que más se asocia a la variable dependiente. Una vez efectuada esta primera segmentación, el proceso se reinicia en cada uno de los subgrupos y

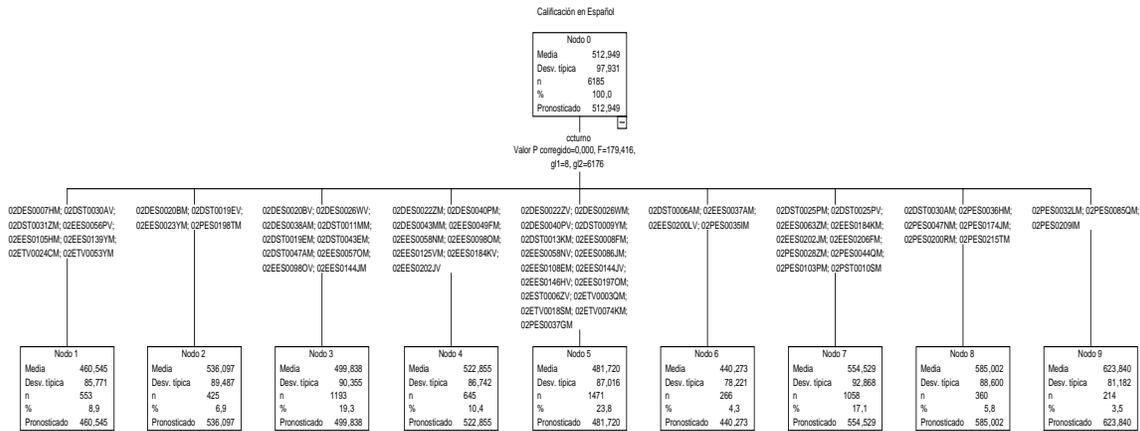
finaliza cuando se alcanza alguno de los criterios de parada establecidos *a priori*. El resultado de un proceso de tales características se plasma en un árbol de decisión que muestra la estructura y relaciones de las variables para cada uno de los segmentos o subgrupos.

Existen varios algoritmos o procedimientos de generación de árboles de decisión. En el presente estudio se empleó el método CHAID (CHI-squared Automatic Interaction Detector), desarrollado por Kass (1980), y en concreto una variante del mismo denominado CHAID exhaustivo, desarrollado por Bigs, de Ville y Suen (1991). La razón básica es que es aplicable a cualquier tipo de variables, tanto dependiente como predictoras y las soluciones que ofrece no son binarias sino que para cada nivel de segmentación, el algoritmo busca la clasificación óptima generando tantos subgrupos como sean necesarios.

En función de los objetivos de este trabajo, este conjunto de técnicas resultan especialmente adecuadas por varias razones: en primer lugar, asumiendo un modelo clásico de dependencia en el que la variable criterio es el rendimiento académico, en este caso en lengua española, estas técnicas nos permiten elaborar modelos con todo tipo de variables predictoras. Segundo, tanto el procedimiento de segmentación como la presentación gráfica de resultados permiten el estudio de relaciones y efectos de interacción entre variables *para subgrupos específicos*, y esto, como inmediatamente vamos a ver es de especial importancia para nosotros. En esta línea, Dudley, Dilorio y Soet (2000) afirman que las técnicas de segmentación, y más específicamente el algoritmo CHAID por el que hemos optado, resultan útiles para detectar y explicar la interacción de datos categóricos, especialmente en el caso de interacciones complejas entre dos o más predictores. En Lizasoain y otros (2003) realizamos una primera aproximación al empleo de estas técnicas estadísticas en la evaluación del rendimiento en lenguas en un estudio evaluativo en la Comunidad Autónoma Vasca.

Al realizar el análisis de segmentación habiendo definido como variable criterio la calificación en la prueba de ENLACE en español y como predictoras *todas* las demás, obtenemos un árbol de 9 nodos siendo la primera variable de segmentación la variable de *identificación de centro-turno* a la que antes se ha hecho referencia. En la figura 1 se muestra la apariencia de este primer nivel del árbol de decisión con los 9 nodos, los 9 subgrupos de centros escolares. Por evidentes razones de espacio, se ha reducido el tamaño del gráfico y no son legibles las etiquetas que identifican los centros escolares que conforman cada nodo.

FIGURA 1. PRIMER NIVEL DE SEGMENTACIÓN SEGÚN CENTRO-TURNO



El nodo cero, el nodo raíz, está formado por toda la muestra: por los 6185 estudiantes de los 71 centros (considerando el turno) cuya media en español es de 512.95. Para cada uno de los 9 nodos del primer nivel, el gráfico proporciona el número de casos, la media y la desviación típica en lengua española de cada subgrupo. En la tabla 4 se ha resumido y ordenado esta información añadiendo además los parámetros del nivel socioeconómico (NSE) medio de los centros escolares de cada subgrupo.

TABLA 4. CARACTERÍSTICAS DE LOS 9 NODOS INICIALES

Nodo CHAID exhaustivo	Número de estudiantes	Número de centros	Media Lengua Española	Desv. típica Lengua Española	Media NSE medio	Desv. típica NSE medio
1	266	4	440,27	78,22	-0,4532	0,62
2	553	8	460,54	85,77	-0,2963	0,42
3	1471	17	481,72	87,01	-0,3208	0,48
4	1193	10	499,84	90,35	-0,2059	0,24
5	645	9	522,85	86,74	-0,4313	0,29
6	425	4	536,09	89,48	0,2541	0,20
7	1058	10	554,53	92,87	0,4527	0,42
8	360	6	585,00	88,60	0,8209	0,17
9	214	3	623,84	81,18	1,4597	0,17
Total	6185	71	512,95	97,93	-,0138	0,63

Antes de comentar los datos de esta tabla y en general este primer resultado, es preciso hacer unas consideraciones previas sobre el procedimiento. Las técnicas de segmentación proceden analizando todas las variables predictoras con el objetivo de determinar cuál de ellas -y qué partición de la misma- es la que maximiza la diferencia entre los grupos con respecto a la variable criterio. Pues bien, en este caso ésta es la variable identificadora *centro-turno*. Esto quiere decir que de entre todas las variables predictoras, es decir, todas las del archivo menos la dependiente, ésta es la que genera una mayor diferencia entre los grupos con respecto al rendimiento en español. Es la que más se asocia. Ninguna otra

agrupación atendiendo a los valores de otra variable predictora es capaz de generar una diferencia mayor, de obtener un mayor valor de F.

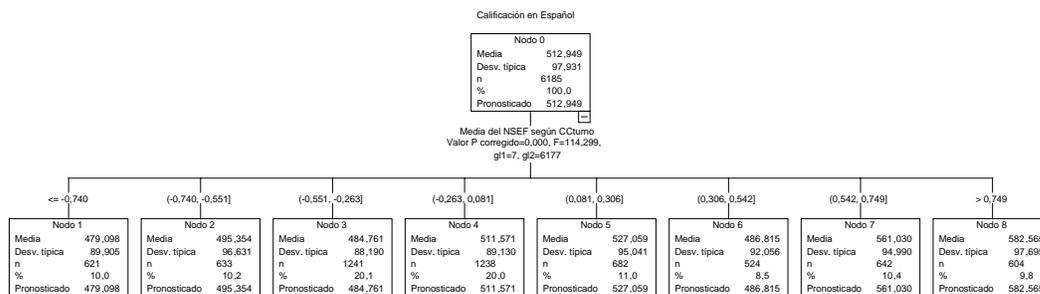
Esto supone que, de cara a explicar las diferencias existentes en el rendimiento en lengua española, el centro escolar es clave. Es el elemento que más contribuye a diferenciar niveles de rendimiento. Hasta aquí no hay nada nuevo pues la importancia del efecto del centro escolar está ampliamente documentada en la literatura.

Como antes se ha comentado, se midieron e incorporaron al estudio numerosas variables del nivel del centro escolar, pero el hecho de que –de todas ellas- sea este identificador la primera variable que segmenta indica que el centro escolar como unidad funcional y organizativa es el elemento clave por encima de cualquiera de los aspectos parciales del mismo tales como la modalidad pública o privada, el clima escolar, la plantilla docente, el nivel socioeconómico medio, etc.

Por tanto, estos 9 nodos constituyen la clasificación óptima de los 71 centros en función de la puntuación en lengua. Pero una vez dicho esto, cabe preguntarse cuáles son los factores de centro que más contribuyen a explicar esta clasificación. Una posible vía para dar respuesta consiste en examinar las características de los colegios de cada uno de los 9 subgrupos. Más adelante haremos una aproximación en esta línea.

Pero las técnicas de segmentación también proporcionan una posible vía de análisis. Como no es muy frecuente que los identificadores operen como variables predictoras, en un segundo proceso el identificador centro-turno se eliminó del modelo y se repitió el análisis obteniendo como resultado que el nivel socioeconómico medio de los centros resultó ser la primera variable de segmentación (figura 2).

FIGURA 2. PRIMER NIVEL DE SEGMENTACIÓN SEGÚN NSE MEDIO UNA VEZ ELIMINADO CENTRO-TURNO



El hecho de que, una vez eliminado del modelo el identificador de centro-turno, sea ésta la primera variable que segmenta indica que, de todas las posibles variables de centro, es el nivel socioeconómico medio de los estudiantes que acuden al centro la variable que más peso tiene. Y eso es fácilmente comprobable si se observa la tabla 4 ó los 8 nodos de la figura 2. Salvo excepciones, como el nodo 5 de la tabla 4 ó el 6 de la figura 2, a mayor NSE medio del centro mayor puntuación media en lengua española. Es una correlación elevada lo que denota que entre todas las variables de centro ésta es la más importante; pero el hecho de que no sea perfecta, de que haya nodos que no se ajustan a la pauta, indica que hay además otros factores que también inciden y explican parte de la varianza del rendimiento.

Las diferencias según el NSE medio son claras y también lo son en función del tipo de escuela (pública/privada) o el turno al que acuden los estudiantes tal y como muestran las tablas 5 y 6 respectivamente en las que aparecen las frecuencias observadas y esperadas.

TABLA 5. TABLA DE CONTINGENCIA DE LOS 9 NODOS Y LA TITULARIDAD DE LOS CENTROS

Titularidad		1	2	3	4	5	6	7	8	9
Privada	Observada	1	13	1	0	18	48	180	138	214
	Esperada	54,8	42,1	118,2	63,9	145,8	26,4	104,9	35,7	21,2
Pública	Observada	552	412	1192	645	1453	218	878	222	0
	Esperada	498,2	382,9	1074,8	581,1	1325,2	239,6	953,1	324,3	192,8

TABLA 6. TABLA DE CONTINGENCIA DE LOS 9 NODOS Y EL TURNO DE LOS CENTROS

Turno		1	2	3	4	5	6	7	8	9
Matutino	Observada	311	352	943	504	933	188	917	360	214
	Esperada	422,2	324,5	910,8	492,4	1123,0	203,1	807,7	274,8	163,4
Vespertino	Observada	242	73	250	141	538	78	141	0	0
	Esperada	130,8	100,5	282,2	152,6	348,0	62,9	250,3	85,2	50,6

Como se puede apreciar, los cuatro primeros nodos, los que agrupan a estudiantes de más bajo rendimiento, se caracterizan por tener un NSE medio bajo, ser mayoritariamente centros públicos y con una mayor proporción de centros vespertinos. Por el contrario, los cuatro nodos de más altas puntuaciones en español concentran mayores proporciones de centros privados en turno matutino y con un NSE medio mayor. Estas tendencias se acentúan cuanto mayor (o menor) sea la puntuación en español de forma que prácticamente no hay centros vespertinos en los nodos altos ni estudiantes que acuden a centros privados en los nodos bajos (ver casillas sombreadas). El nodo central, el quinto, se aparta de la tendencia general puesto que con un NSE medio bajo (el segundo más bajo), obtiene una puntuación media en rendimiento. Los sujetos y centros que conforman este grupo podrían ser ejemplos de resiliencia académica aunque para probar esta hipótesis se requiere un estudio pormenorizado de los mismos que excede de los objetivos de este trabajo.

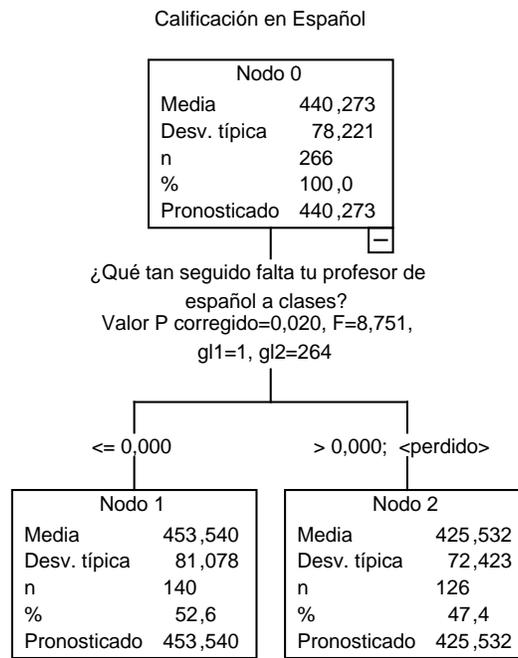
Las técnicas de segmentación no sólo proporcionan una clasificación de los centros escolares en función del rendimiento. Si se opta por que el algoritmo continúe desarrollando el árbol de clasificación, nos encontramos con un modelo de clasificación y regresión que permite observar el efecto de las variables predictoras sobre el rendimiento en lengua española.

Así, en el modelo elaborado por CHAID exhaustivo, en el segundo nivel, y para cada uno de los 9 nodos iniciales, se observa que las variables de segmentación son reactivos concretos pertenecientes a constructos como los siguientes: clima escolar (razones para ir a secundaria, me siento incómodo en mi escuela), nivel socioeconómico o cultural de la familia, gusto por la lectura, motivación y apoyo familiar, puntualidad del profesor, etc. (en esta ocasión el árbol es demasiado complejo y de difícil apreciación por lo que se ha optado por describirlo).

En resumen, el modelo propuesto incorpora las variables que se suelen considerar predictores habituales del rendimiento académico. Algunas de estas variables aparecen con una cierta frecuencia, en diversos niveles y en todas las regiones del árbol (como es el caso del NSE medio de los centros, reactivos del

clima escolar o algunos indicadores de NSE o cultural familiar). Esto es un indicador de la importancia relativa de las mismas dentro del modelo explicativo.

FIGURA 3. SEGUNDO NIVEL DE SEGMENTACIÓN PARA EL NODO 1



Pero hay otro conjunto de variables que aparecen *sólo en algunos niveles* más profundos del árbol o en algunas regiones. Por ejemplo, la figura 3 muestra el segundo nivel de segmentación para el nodo 1 que agrupa los 266 casos que obtienen la menor puntuación en lengua (440.273). Aquí la primera variable de segmentación no es ni el clima escolar ni ninguna de las variables antes apuntadas y que son habituales en modelos explicativos. Se trata de una variable ligada específicamente al cumplimiento formal del horario por parte del profesor. Como puede observarse, en aquellos casos en que el docente falta habitualmente a clase la media desciende hasta los 425,532 puntos.

Este es un ejemplo del comportamiento de este tipo de variables en estos casos. Su capacidad de segmentación opera exclusivamente en algunos subgrupos específicos de sujetos. Y lo que aquí aparece es que para grupos pequeños pero muy diferenciados con respecto a la puntuación en lengua española, aparecen variables predictoras muy distintas. En algunos casos (de rendimiento bajo) aparecen variables ligadas con el cumplimiento formal del docente o del estudiante (puntualidad, falta de asistencia) o el estado físico de las aulas o los sanitarios. En otros surgen (y con relación inversa) la existencia o uso de determinadas tecnologías educativas.

Tal y como se apuntaba anteriormente, en el sistema educativo de Baja California conviven realidades sociales, culturales y económicas muy heterogéneas. Esto hace que las diferencias entre centros escolares sean muy acusadas de forma que junto a centros escolares de muy alto nivel nos encontramos con otros que se encuentran en el otro extremo de la escala. Esto hace que sea difícil encontrar y ajustar un único modelo para realidades sociales tan dispares, pues se observan dos conjuntos de indicadores o predictores que operan de forma distinta en diferentes subgrupos de centros:

- En primer lugar, hay variables con gran capacidad de incidencia en el rendimiento académico que aparecen como variables de segmentación de forma repetida en numerosos niveles y nodos. Tal es el caso, por ejemplo, del nivel socioeconómico o cultural o el clima escolar.
- Pero junto a éstas, hay un segundo conjunto de variables que operan de forma casi exclusiva en algunos subgrupos de algunos niveles, habitualmente en los nodos extremos. Y así, en los nodos altos surgen cuestiones propias de sistemas educativos avanzados (uso de recursos tecnológicos, conexiones en el aula a Internet, etc.), mientras que en los bajos aparecen otras que se señalan como oportunidades básicas y que están más ligadas al cumplimiento efectivo del currículum, a la asistencia y puntualidad de profesores y alumnos o al estado físico de las instalaciones escolares.

2.4.2. Modelización mediante regresión multinivel

Consecuentemente con lo anteriormente expuesto, a continuación se procede al diseño y ajuste de un modelo explicativo del rendimiento en lengua española. Dado que los datos presentan una estructura anidada (estudiantes en centros escolares), y vista además la importancia que el *centro-turno* tiene como primera variable de segmentación, parece claro que para este fin se ha de realizar una aproximación multinivel mediante modelos jerárquicos lineales.

Una de las utilidades más relevantes de los árboles de decisión es el uso de los mismos para seleccionar variables relevantes para modelos predictivos más rigurosos por lo que, en aplicación de la misma, se consideran como variables candidatas a predictores aquellas que en los árboles han mostrado una alta capacidad discriminativa bien porque hayan aparecido en niveles iniciales de los árboles, bien porque aparezcan con frecuencia para determinados nodos en diversos niveles.

Posteriormente, y de cara a un análisis diferencial de los subgrupos más extremos, se tomarán en consideración también aquellas variables que sólo hayan aparecido en algunos niveles de segmentación de estos nodos.

En consecuencia, para el diseño y ajuste de un modelo multinivel de toda la muestra las variables consideradas han sido:

Variable dependiente: calificación en la prueba de ENLACE de Lengua Española. Es una puntuación normalizada de media 500 y desviación típica 100.

Variabes predictoras del primer nivel (estudiante):

- Género de los estudiantes codificada como 0 (cero) para los estudiantes masculinos y 1 (uno) para las chicas.
- Trabajo remunerado: 0 no trabaja, 1 sí.
- Gusto por la lectura: 0, no gusta; 1, gusto moderado; 2, gusto acusado.
- Rendimiento previo en lengua española medido por las calificaciones en esta materia del curso anterior en una escala que va de 0 a 4.
- Nivel socioeconómico familiar. Puntuación resultante de un análisis de componentes principales (ACP) realizado sobre los reactivos relativos a indicadores sociales, económicos y culturales de la unidad familiar. Es una escala tipificada con media 0 y desviación típica 1.

- Clima escolar. Puntuación resultante de un ACP efectuado sobre los reactivos relativos al clima en el aula y en el centro, la relación con los compañeros y los docentes. Es una escala tipificada con media 0 y desviación típica 1.

Adicionalmente, y de cara al estudio diferencial de los subgrupos extremos se consideraron:

- Puntualidad del profesor de español. Juicio que el estudiante expresa sobre el nivel de puntualidad del docente en una escala que va desde 0 (nada puntual) a 3 (muy puntual).
- Faltas de asistencia del profesor de español. Juicio que el estudiante expresa sobre las faltas de asistencia a clase del docente en una escala que va desde 0 (nunca falta) a 3 (siempre o casi siempre falta). Nótese que estas dos variables relativas al cumplimiento formal del docente tienen escalas opuestas, inversa la puntualidad, positiva la del nivel de asistencia.

Variables predictoras del segundo nivel (*centro-turno*):

- Turno del centro escolar: 0, matutino; 1, vespertino.
- Titularidad del centro escolar: 0, privada; 1 pública.
- Nivel socioeconómico medio del centro. Promedio para cada centro-turno del nivel socioeconómico familiar de los estudiantes de dicho centro.
- Clima escolar medio del centro. Promedio para cada centro-turno del clima escolar percibido por los estudiantes de dicho centro.

Al igual que antes, para el estudio de los subgrupos extremos:

- Puntualidad media de los profesores de español. Promedio para cada *centro-turno* de la puntuación asignada por los estudiantes de dicho centro al nivel de puntualidad de su docente.
- Nivel medio de asistencia de los profesores de español. Promedio para cada *centro-turno* de la puntuación asignada por los estudiantes de dicho centro al nivel de asistencia de su docente.

Lógicamente, en estas variables agregadas del segundo nivel, la escala es la misma que en las correspondientes del primer nivel.

Modelo HLM

Mediante el proceso habitual y empleando el programa HLM, se procede a elaborar y validar un modelo jerárquico. Dicho modelo incluye los efectos significativos, es decir, se enfoca el problema desde un punto de vista exploratorio y no buscando el modelo más parsimonioso.

El resultado es el siguiente:

$$\begin{aligned} \text{Puntuación en español} = & 443,21 + \\ & + 13,53 * \text{Sexo (varón} = 0) + \\ & - 16,17 * \text{Trabaja (no} = 0) + \\ & + 30,27 * \text{Gusto por la lectura (no} = 0, \text{poco} = 1, \text{mucho} = 2) + \\ & + 14,03 * \text{Calificación en español en 2}^\circ \text{ Primaria (0 a 4) +} \end{aligned}$$

- + 4,43 * NSE Familiar (-3,26 a 2,69) +
- + 39,20 * NSE Escuela (-1,76 a 1,53) +
- + 28,83 * Clima escolar (-0,76 a 1,34)

Parte aleatoria:

Vaarianza entre escuelas:	682,46
Varianza intra-escuelas :	6919,98

Controlando la calificación en *Lengua española* en 2º de Secundaria, los resultados del modelo indican que la puntuación media en *Lengua española* de los estudiantes varones de 3º de Secundaria que *no trabajan y no les gusta leer*, cuyo *NSE familiar* y *escolar* son medios, al igual que *el clima escolar*, es de algo más de 440 puntos (recuérdese que la media es aproximadamente 500 y la desviación típica 100). A ello hay que añadir algo más de 13 puntos a las alumnas; además, a los estudiantes que trabajan hay que restarles algo más de 16 puntos y añadir más de 60 puntos a los estudiantes a los que *gusta la lectura*. Por otra parte por cada punto de incremento del *NSE familiar*, el incremento en la puntuación en *Lengua española* resulta ser de casi 4,5 puntos (téngase presente que los valores del NSE Familiar oscilan entre -3 y 3 aproximadamente); el efecto del *NSE de la Escuela* (éste oscila en punto y medio alrededor de la media) es más acusado, pues ahora cada punto de incremento en él produce un incremento de casi 40 puntos en la puntuación en Lengua española. Por su parte, 1 punto de incremento en el *Clima escolar* global de la escuela (que escila alrededor de la media en un punto aproximadamente) origina un incremento de casi 30 en Lengua.

La parte aleatoria del modelo consiste en que la varianza entre escuelas es de 682, 46 y la varianza intra-escuelas, 6919,98; por lo tanto la correlación intra-clases es del 9% aproximadamente.

Los resultados obtenidos en este proceso de modelización, nos muestran un conjunto de predictores y una ecuación multinivel que, *grosso modo*, no puede considerarse muy distinta de los resultados obtenidos en estudios similares.

En el primer nivel, el género de los estudiantes muestra el habitual efecto positivo para las chicas y el efecto negativo de simultanear estudios y trabajo no merece comentario adicional. Es igualmente sabido que el rendimiento académico previo es un buen predictor del rendimiento actual y lo mismo ocurre con el nivel socioeconómico familiar. Por último, el gusto por la lectura es un importante factor de éxito en una materia como la lengua española y, aunque las escalas no sean directamente comparables, su efecto se puede equiparar al del nivel socioeconómico familiar. Resultados similares encontramos en un modelo jerárquico realizado para el estudio longitudinal de los efectos de modelos lingüísticos propios del sistema educativo vasco sobre el rendimiento en español (Santiago y otros, 2008).

En el nivel de los centros escolares nos encontramos sólo con dos variables: la media del nivel socioeconómico y la media del clima escolar. Ambas tienen un importante efecto directo sobre el rendimiento en español, especialmente el nivel socioeconómico medio. A este respecto, se debe reconocer el impacto de las diferencias socioeconómicas y culturales en el rendimiento de los estudiantes de Baja California y del propio sistema educativo mexicano al igual que ocurre en otros países y sistemas educativos. De la misma manera, se confirma también el *efecto contextual* (Willms, 1986; Rumberger & Willms, 1992; Willms, 2003; Willms, 2004; Tajalli & Opheim, 2004; Howley & Howley, 2004; Warschauer

et al., 2004; Willms, 2006; Lizasoain y otros, 2007) de forma que el nivel socioeconómico medio (NSE) de las escuelas tiene mucho peso a la hora de explicar las diferencias de rendimiento.

Puede llamar la atención el hecho de que no aparezca la titularidad del centro o el turno, aunque es altamente probable que los efectos de estas variables se hayan subsumido en el del nivel socioeconómico medio. Por último, las variables relativas al cumplimiento formal de los docentes tanto del nivel 1 como del 2, no han resultado significativas.

Estudio diferencial

Una vez elaborado y ajustado un modelo para la totalidad de la muestra, se aborda el estudio diferencial de los subgrupos extremos. La razón es, como antes se ha visto, que en el proceso de segmentación en los nodos de estos subgrupos operan como variables de alta capacidad discriminativa algunas que no aparecen en otros lugares del árbol. Con objeto de realizar una primera aproximación a este enfoque diferencial, se seleccionan los 480 casos pertenecientes a los nodos 1 y 9 que, como se ha dicho, tienen respectivamente las puntuaciones medias más baja y alta en lengua española.

El nodo 1 está compuesto por los 266 estudiantes que, agrupados en 4 centros, obtienen le promedio más bajo en lengua española (440,27); mientras que el nodo 9 lo componen los 214 estudiantes pertenecientes a 3 centros y obtienen un promedio de 623,84.

En las tablas 7 y 8 se caracterizan ambos subgrupos en función de distintas variables. Es reseñable el hecho de que todos los centros del nodo 9 sean de titularidad privada y que ninguno imparta docencia en turno vespertino. Igualmente se observan otras diferencias importantes: en el nodo 1 hay mayor proporción de chicos que de chicas lo que probablemente se deba a una mayor presencia de estudiantes repetidores entre los que los chicos son mayoritarios. También es acusada la diferente proporción entre los que declaran trabajar en el nodo 1 (68, casi el 30%) frente a los 18 (casi el 9%) que están en la misma situación en el nodo 9. Por último las diferencias en el nivel de gusto por la lectura, en el clima escolar, en el NSE y en el cumplimiento formal son siempre en sentido favorable al nodo 9. Todas estas diferencias resultan significativas con una $p < 0,01$.

TABLA 7. CARACTERIZACIÓN DE LOS NODOS 1 Y 9 (I)

Nodos (%)	Género		Trabajo		Gusto lectura			Titularidad		Turno	
	H	M	No	Sí	No gusta	Poco	Mucho	Privada	Pública	Matut.	Vesp.
1	57,4	42,6	70,8	29,2	22,6	65,5	11,9	18	82	70,7	29,3
9	50,7	49,3	91,2	8,8	17,7	56,5	25,8	100	0	100	0

TABLA 8. CARACTERIZACIÓN DE LOS NODOS 1 Y 9 (II)

Nodos	Número de estudiantes	Número de centros	Media Lengua Española	Media NSE	Media Clima escolar	Puntualidad	Asistencia
1	266	4	440,27	-0,4532	-0,0791	1,95	0,57
9	214	3	623,84	1,4597	0,8229	2,32	0,18

Pues bien, a la vista de estas diferencias y de la presencia en estos nodos de nuevas variables relevantes, ahora se trata de probar la hipótesis inicialmente apuntada de que los modelos explicativos del

rendimiento en lengua española serán distintos para estos grupos con respecto al elaborado para el total de la muestra.

Para ello pueden plantearse varias vías analíticas. Una primera consistiría en segmentar el archivo de datos en función del nodo tratando y estudiando cada nodo por separado formulando y ajustando un modelo para cada uno. Pero esto no es una aproximación adecuada porque estos nodos no constituyen una agrupación natural de las unidades sino que se han establecido -mediante los árboles de decisión- en función precisamente de las puntuaciones en rendimiento en lengua española. Esto supondría realizar un análisis circular y cerrado: la variable dependiente es español y los subgrupos están definidos en función de ella. Además, al tratarse de agrupaciones de centros es lógico pensar que la varianza entre ellos sea menor lo que dificultaría el ajuste de modelos multinivel.

Una posible alternativa es centrar el análisis diferencial estudiando una submuestra compuesta sólo por los dos subgrupos extremos, el 1 y el 9. Al retirar los subgrupos centrales, mediante esta selección nos quedamos sólo con los estudiantes y centros que obtienen puntuaciones más extremas. Aquí sí cabe plantearse un modelo jerárquico lineal, pero en este caso hay una dificultad que radica en que se cuenta con muy pocas unidades del nivel 2. Recordemos que en este caso sólo hay 4 centros en el nodo 1 y 3 en el 2 lo que arroja un total de 7 que es claramente insuficiente para realizar modelización multinivel.

Esto origina que en HLM-5 los modelos multinivel no se desarrollen de forma adecuada pues en los resultados no se muestra el valor de la significación por problemas con los grados de libertad. Ello puede ser debido a este reducido número de elementos del segundo nivel. Así el modelo resulta inestable y además poco convincente (Gaviria y Castro, 2005). O, como afirman Bower y Drake (2005), las estimaciones con menos de 10 grados de libertad en el nivel 2 pueden ser inconsistentes.

En consecuencia, se ha procedido a obtener el modelo sin componentes aleatorios, es decir, aplicando una regresión lineal múltiple. Y los resultados son:

$$\begin{aligned} \text{Puntuación en español} = & 480,66 + \\ & + 15,11 * \text{Calificación en español en 2º Primaria (0 a 4)} + \\ & + 23,67 * \text{Gusto por la lectura (no = 0, poco = 1, mucho = 2)} + \\ & + 29,27 * \text{Clima escolar (-3 a 2)} + \\ & + 17,69 * \text{NSE Familiar (-3 a 2,5)} - \\ & - 83,14 * \text{Red (pública = 1)} - \\ & - 11,78 * \text{Faltas de asistencia del profesor (nunca = 0, a siempre = 3)} \end{aligned}$$

El valor de R^2 corregida es del 51% y la varianza de los residuos $\sigma^2 = 7178,60$.

Por lo tanto, controlando la calificación en *Lengua española* en 2º de Secundaria, los resultados del modelo indican que la puntuación media en *Lengua española* de los estudiantes de 3º de Secundaria de *centros privados* a los que no les *gusta la lectura*, cuyo profesor no falta nunca o casi nunca a clase, cuyo *NSE familiar* es medio, al igual que el *clima escolar*, es de algo más de 470 puntos (la media es aproximadamente 520 y la desviación típica 120). A ello hay que añadir casi 50 puntos a los estudiantes a los que *gusta la lectura*. Por su parte, 1 punto de incremento en el *Clima escolar* según el estudiante (que oscila entre -3 y 2 puntos) origina un incremento de algo más de 30 en *Lengua*. Por otra parte por cada punto de incremento del *NSE familiar*, el incremento en la puntuación en *Lengua española* resulta

ser de casi 19 puntos (téngase presente que los valores del NSE Familiar oscilan entre -3 y 2,5 aproximadamente). Estudiar en centros públicos tiene un efecto de disminución en la puntuación en *Lengua española* de más de 80 puntos en comparación con los privados y 1 punto de incremento en la escala de falta de asistencia del docente conlleva una disminución de casi 12 puntos en el puntaje de lengua.

En este modelo vemos que las variables relevantes son muy similares a las del modelo multinivel para toda la muestra pero ello no es óbice para que entre ambos modelos no deje de haber diferencias reseñables.

Comenzando por las similitudes, se observa que ambos modelos comparten 4 predictores: el gusto por la lectura, el rendimiento previo, el nivel socioeconómico y el clima escolar. A la vista de esto parece clara la relevancia de las cuatro para explicar las diferencias en el desempeño en lengua española. Aquí se han incluido obviamente las variables correspondientes al nivel 1 de los estudiantes.

Pero junto a estas similitudes, también hay algunas diferencias a tomar en consideración. En primer lugar, el sexo de los estudiantes aquí no resulta significativo a pesar de que como se ha comentado en lo relativo al rendimiento en lengua la ventaja de las chicas está ampliamente documentada. En cambio, aparece una variable relativa al cumplimiento formal de los docentes cual es la de la frecuencia de sus faltas de asistencia. Cada incremento de un punto en la escala (que va de 0 a 3) lleva aparejada una disminución de casi 12 puntos en el rendimiento en español. Aunque su nivel de significación está justo en el límite de lo admisible ($p=0,05$) se incorpora al modelo refrendando lo que ya habían apuntado los modelos de los árboles de decisión, a saber, en los centros extremos los niveles de cumplimiento formal por parte de los docentes tienden también a ser extremos: muy altos en los centros de medias elevadas, muy bajos en los centros de bajo rendimiento.

Estos resultados han de ser tenidos en cuenta con precaución habida cuenta de la imposibilidad de elaborar y validar modelos multinivel para estos subgrupos extremos. Pero parece que existe alguna tendencia a confirmar la hipótesis formulada en el sentido de que en sistemas educativos poco homogéneos conviven realidades sociales muy dispares y, por tanto, los subgrupos extremos tienen un comportamiento diferente al del gran grupo. Sea como sea, parece evidente que es necesario realizar estudios con muestras con mayor número de elementos del nivel 2.

3. CONCLUSIONES

En primer lugar, y tal y como ya se ha señalado, el centro escolar (considerando el turno) como unidad funcional es el elemento clave a la hora de establecer diferencias en el rendimiento en lengua española de los estudiantes. En esta línea, de todas las variables de centro que pueden contribuir a explicar esta capacidad discriminativa, es el nivel socioeconómico medio de los estudiantes que acuden a la escuela la variable que más peso tiene.

Para el caso de la población involucrada en el presente estudio, las variables del turno y tipo de escuela son determinantes en la explicación del rendimiento académico, de forma que los estudiantes de centros privados tienden a obtener un mejor desempeño escolar mientras que los estudiantes que acuden al turno vespertino obtienen los resultados académicos más pobres. Ambas variables son concomitantes con el NSE medio pues los centros privados tienen un NSE medio mayor, mientras que los estudiantes que acuden a realizar sus estudios en horario vespertino presentan un NSE medio inferior.

Continuando con las cuestiones de naturaleza socioeconómica, en la población en edad escolar en Baja California hay un porcentaje importante de estudiantes que declaran trabajar (el 18% en nuestra muestra). Esto se asocia negativamente con el rendimiento académico de forma que en aquellos subgrupos de menor rendimiento la proporción de estudiantes que trabajan es significativamente mayor.

En lo relativo a aspectos de naturaleza escolar o vinculados con la propia actividad educativa, se cuenta con evidencia que apunta en la misma dirección que los resultados de diversos estudios con propósitos similares, de forma tal que variables como el gusto por la lectura o un clima escolar favorable, se asocian con calificaciones altas en lengua española.

Por último, en lo relativo al cumplimiento formal de docentes y estudiantes (faltas de asistencia, puntualidad, etc.), éstas resultan tener una alta capacidad explicativa pero sólo para algunos subgrupos con estudiantes con calificaciones bajas o muy bajas.

En el sistema educativo de Baja California conviven realidades sociales, culturales y económicas muy heterogéneas. Esto hace que las diferencias entre centros escolares sean muy acusadas de forma que junto a centros escolares de muy alto nivel nos encontramos con otros que se encuentran en el otro extremo de la escala. Esto dificulta el diseño y ajuste de un único modelo para analizar realidades sociales tan dispares, pues junto con los indicadores y predictores que tienen una alta capacidad explicativa en todos los subgrupos y niveles, nos encontramos con otras variables que tienen también una alta incidencia y capacidad explicativa pero sólo en algunos subgrupos de algunos niveles, habitualmente en los nodos extremos.

En las escuelas de muy bajo nivel socioeconómico medio, las oportunidades de aprendizaje de sus estudiantes no radican tanto en disponer de ordenadores o en estar conectados a la red. El asunto es más sencillo y elemental: como primera condición basta con que los profesores y estudiantes acudan a clase con regularidad y puntualidad llevando a cabo unos y otros las tareas encomendadas tendentes al cumplimiento efectivo del currículo.

Baja California, y nos atreveríamos a decir que México en su conjunto es un país complejo con una realidad social también compleja que hace que coexistan centros perfectamente equiparables a los de nivel socioeconómico alto de los países desarrollados, con otras situaciones y centros propios de países en vías de desarrollo.

En nuestra opinión, tanto desde un punto de vista centrado en la elaboración y ajuste de modelos explicativos del rendimiento académico como en aquel que priorice el diseño y puesta en marcha de acciones y programas de mejora es muy importante tomar en consideración este efecto diferencial. Es preciso considerar estas variables que operan como oportunidades de aprendizaje sólo en algunos subgrupos aunque sean pequeños y extremos, porque en caso contrario para estos sectores de la población (que son los más desfavorecidos) los modelos generales y las acciones de mejora que eventualmente se puedan diseñar y en su caso aplicar estarán desenfocados y por tanto no resultarán de utilidad.

REFERENCIAS BIBLIOGRÁFICAS

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. (2008). El aprendizaje en tercero de preescolar en México. México: INEE.
- Barrera, O. (2009). La Evaluación de la Educación Cívica y Ética en México. Un Recorrido por 3º de Primaria y 3º de Secundaria. *Revista Iberoamericana de Evaluación Educativa*, 2 (1), pp.114-129. Consultado el 10 de octubre de 2010 en <http://www.rinace.net/riee/numeros/vol2-num1/art6.pdf>.
- Biggs, D.B. de Ville and Suen, E. (1991). A method of choosing multiway partitions for classification and decision trees. *Journal of Applied Statistics*. 18, pp. 49-62.
- Bowers, J., Drake, K. W. (2005). EDA for HLM: Visualization when Probabilistic Inference Fails. *Political Analysis*, 13(4), pp. 301-326.
- Carvalho, M., Caso, J. and 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 8 de abril de 2010 en <http://redie.uabc.mx/vol9no2/contenido-carvalho.html>.
- Caso, J., Hernández, E. y Rodríguez, J.C. (2009). Análisis Multinivel de los Factores Asociados a la Educación Cívica en Estudiantes de Secundaria. *Revista Iberoamericana de Evaluación Educativa*, 2 (1), pp. 226-238. Consultado el (Fecha) en <http://www.rinace.net/riee/numeros/vol2-num1/art12.pdf>.
- Dudley, W., Dilorio, C., Soet, J. (2000). Detecting and explicating interactions in categorical data. *Nursing-Research*, 49(1), pp. 53-56.
- Gaviria, J.L., Castro, M. (2005). Modelos jerárquicos lineales. Madrid: Ed. La Muralla.
- Howley, C. B. and Howley, A.A. (2004). School size and the influence of socioeconomic status on student achievement: Confronting the threat of size bias in national data sets. *Education Policy Analysis Archives*, 12(52). Recuperado de <http://epaa.asu.edu/epaa/v12n52/>.
- Kass, G. (1980). An exploratory technique for investigating large quantities of categorical data. *Applied Statistics*. 29(2), pp. 119-127.
- Lizasoain, L., Joaristi, L., Lucas, J.F., Santiago, C. (2007). El efecto contextual del nivel socioeconómico sobre el rendimiento académico en la Educación Secundaria Obligatoria. Estudio diferencial del nivel socioeconómico familiar y del centro. *Education Policy Analysis Archives*, 15(6), pp. 1-25.
- Lizasoain, L., Santiago, C., Moyano, N., Munarriz, B., Joaristi, L., Lucas, J.F., Sedano, M. (2003). El uso de Las Técnicas de Segmentación en la evaluación del rendimiento de Lenguas. Un estudio en la Comunidad Autónoma Vasca. *Revista de Investigación Educativa*, 21(1), pp. 93- 111.
- Ruíz, G. (2009). La Calidad del Sistema Educativo Mexicano desde los Resultados de Evaluaciones Nacionales. El Aprendizaje en Matemáticas. *Revista Iberoamericana de Evaluación Educativa*, 2 (1), pp. 74-89. Consultado el 10 de octubre de 2010 en <http://www.rinace.net/riee/numeros/vol2-num1/art4.pdf>.
- Rumberger, R., and Willms, J.D. (1992). The impact of racial and ethnic segregation on the achievement gap in California high schools. *Educational Evaluation and Policy Analysis*, 14(4), pp. 337-396.
- Sánchez, A. (2009). La Evaluación del Logro Escolar en la Educación Básica en México. Los Excale de Matemáticas. *Revista Iberoamericana de Evaluación Educativa*, 2(1), pp. 186-204. Consultado el 10 de octubre de 2010 en <http://www.rinace.net/riee/numeros/vol2-num1/art10.pdf>.

- Santiago, C., Lucas, J.F., Joaristi, L., Lizasoain, L., Moyano, N. (2008). A longitudinal Study of Academic Achievement in Spanish: The Effect of Linguistic Models. *Language, Culture And Curriculum*, 21(1), pp. 48-58.
- Tajalli, H. and Opheim, C. (2004). Strategies for Closing the Gap: Predicting Student Performance in Economically Disadvantaged Schools. *Educational Research Quarterly* 28(4), pp.44-54.
- Valenti, G. (2007) Factores asociados al logro educativo de matemáticas y español en la Prueba ENLACE 2007. México: SEP.
- Warschauer, M., Knobel, M. and Stone, L. (2004). Technology and Equity in Schooling: Deconstructing the Digital Divide. *Educational Policy*, 18(4), pp. 562-588.
- Willms, J.D. (1986). Social Class Segregation and Its Relationship to Pupil's Examination Results in Scotland. *American Educational Review* 51(2), pp.224-241.
- Willms, J.D. (2003). *Ten Hypotheses about Socioeconomic Gradients and Community Differences in Children's Developmental Outcomes*. Final Report. Applied Research Branch. Strategic Policy. Human Resources Development. Canada. Consultado el 26 de abril en <http://www.hrdc-drhc.gc.ca/sp-ps/arb-dgra>.
- Willms, J.D. (2004). *Reading Achievement in Canada and the United States: Findings from the OECD Programme of International Student Assessment*. Final Report. Learning Policy Directorate Strategic Policy and Planning Human Resources and Skills Development. Canada. Consultado el 25 de abril en <http://www.hrsdc-rhdcc.gc.ca/sp-ps/arb-dgra>.
- Willms, J.D. (2006). *Learning Divides: Ten Policy Questions About The Performance And Equity Of Schools And Schooling Systems*. Montreal: UNESCO Institute for Statistics.



PROBLEMAS DE EQUIDAD EN EL SISTEMA EDUCATIVO HONDUREÑO

EQUITY PROBLEM IN THE HONDURAN EDUCATIVE SYSTEM

Mario Alas Solís y Germán E. Moncada Godoy

Revista Iberoamericana de Evaluación Educativa 2010 -Volumen 3, Número 3

<http://www.rinace.net/riee/numeros/vol3-num3/art7.pdf>

Fecha de recepción: 16 septiembre de 2010
Fecha de dictaminación: 24 octubre de 2010
Fecha de aceptación: 04 de noviembre de 2010

“La constatación inicial es que las desigualdades escolares no se originan en el sistema educativo sino que se agravan en él.” Álvaro Marchesi (2000).

En la sociedad contemporánea, la educación se consigna como un derecho universal. Para muchas personas, la educación puede ser la única forma de escapar de una vida de pobreza; para las sociedades, es crucial para su desarrollo socioeconómico. Se espera que la educación apoye el desarrollo de valores tales como honestidad, equidad, justicia social, etc., y el sentido de responsabilidad compartida que es clave para el tejido social y las formas de gobierno democráticas. Pero también se espera que desarrolle competencias tanto cognoscitivas como procedimentales y actitudinales, que permitan a los educandos incorporarse exitosamente en las estructuras productivas de la sociedad (Fondo Mundial de Educación, 2000).

Entendiendo que la educación es un factor fundamental para vincular la integración social y el desarrollo económico en el contexto de las sociedades contemporáneas, existe hoy -sobre esta vinculación-consenso en la bibliografía especializada respecto de que: “A la pregunta por el tipo de desarrollo que podemos impulsar, sobreviene sin dilación la pregunta por el tipo de educación de que disponemos. Este nexo causal torna impostergable enfrentar los problemas de **calidad** y **equidad** en la educación” (Hopenhayn, 1998).

En relación con estos problemas, en el ámbito de las políticas educativas existe un largo debate entre expertos y diseñadores de políticas que priorizan el criterio de la equidad, frente a quienes priorizan la calidad. También están quienes intentan realizar síntesis armónicas o suavizar esta disyuntiva. Una de los mayores puntos de controversia se plantea el torno al problema entre el logro de una elevada calidad educativa con el logro de un grado satisfactorio de equidad, o por el contrario, la priorización de la equidad sobre la calidad. Este dilema entre calidad y equidad es muy relevante desde luego, ya que incide directamente en el diseño de políticas de distribución de recursos al interior de los sistemas educativos.

Y la distribución de recursos, al igual que **la equidad en educación hace referencia a los principios básicos de la sociedad democrática**. Desde el nacimiento de los ideales democráticos, la meta de una educación universal e igualitaria ha sido uno de los ideales básicos de una sociedad que se considere justa y democrática. Actualmente, dada la complejidad del funcionamiento de las sociedades, esta importancia se acrecienta, ya que es imposible pensar en un ejercicio informado de la ciudadanía y un control popular de la burocracia gubernamental, sin que esta población haya recibido una cantidad considerable de educación.

Por tanto, la pregunta por lo que es justo/ injusto en las sociedades democráticas contemporáneas lleva al examen de las desigualdades educativas y a plantearse dos interrogantes fundamentales (García-Huidobro, 2003): ¿Cuál es el nivel de igualdad educativa que se debe exigir para la educación obligatoria? ¿Cuándo las desigualdades educativas se pueden considerar “aceptables” en la educación especializada? El concepto de educación común u obligatoria se funda en que provee los elementos culturales necesarios para vivir en condiciones plenas en la sociedad referida. Se trata de que proporcione una formación educativa básica relacionada con la ciudadanía, con la identidad social, con la capacidad de incorporarse a la estructura productiva así como aportar fundamentos para que las personas puedan ejercer la libertad en esas sociedades. En este sentido se espera lograr una cierta igualdad de resultados educativos, entendida como el logro por parte de todos de un conjunto determinado de competencias básicas. Se trata de un criterio de equidad muy exigente, para el cual no es suficiente que los niños y

jóvenes asistan a las escuelas y progresen en ellas, sino que reivindica que todos logren un nivel similar de aprendizajes (García-Huidobro, 2003).

En este marco referencial, los actuales esfuerzos por mejorar la calidad de la educación en América Latina se focalizan en la búsqueda de maneras eficientes y eficaces de gestionar los limitados recursos disponibles para mejorar la equidad en la educación, tratando de reducir las brechas en el desempeño escolar asociadas a las diferencias de estrato socioeconómico y, en particular, a la transmisión intergeneracional de la pobreza (Cohen, E., 2002). Sin embargo, las políticas sociales que abordan el problema de la equidad en educación en América Latina se enfrentan con las dificultades acumuladas por décadas en un contexto de extrema desigualdad social.

1. EQUIDAD EN EDUCACIÓN EN UN CONTEXTO DE PROFUNDAS DESIGUALDADES SOCIALES

Las políticas de equidad educativa en el subcontinente latinoamericano se enfrentan a un contexto particularmente difícil. América Latina no es la región más pobre del mundo pero sí la más desigual. La opulencia y la extrema miseria conviven peligrosamente al interior de estas sociedades: **“América Latina es la región más inequitativa del mundo**, en ella persisten la pobreza y la desigualdad distributiva que han dificultado la constitución de sociedades más integradas. La desigualdad de oportunidades educativas, en función del nivel socioeconómico de los estudiantes, es una de las más agudas en la región, de lo que se deduce que la educación no está siendo capaz de romper el círculo vicioso de la pobreza y la desigualdad social” (Blanco, y Cusato, 2004).

El derecho a la educación todavía no se cumple para toda la población en los países de América Latina. Todavía hay muchos niños y niñas que enfrentan dificultades para su incorporación a los sistemas educativos, ya sea porque están excluidos o porque reciben una educación de menor calidad que provoca que alcancen menores niveles de aprendizaje. La equidad en la educación implica que se cumplan los derechos a la igualdad de oportunidades, y a la no discriminación. En la Convención de los Derechos del Niño se señala, explícitamente, que el derecho a la educación ha de estar eximido de cualquier tipo de discriminación y ha de estar inspirado en la igualdad de oportunidades.

Como han señalado numerosos autores (Sen, 1995; Hopenhayn, 1998; Cohen, 2002), la igualdad de oportunidades en educación implica no sólo igualdad en el acceso, sino y sobretodo, igualdad en la calidad de la educación que se brinda y en los logros de aprendizaje que alcanzan los alumnos. **Es claro que la igualdad de acceso no es suficiente para garantizar la equidad educativa si no se proporcionan programas de similar calidad** a todos los niños y niñas, apoyándoles para lograr aprendizajes semejantes, sea cual fuere su condición social, cultural e individual. Para lograrlo los gobiernos y sistemas educativos abordan el problema de proveer los recursos humanos, materiales y financieros necesarios para que todos los alumnos cuenten con las oportunidades adecuadas para alcanzar los niveles de aprendizaje esperados. Pero esta aspiración se confronta con contextos sociales profundamente desiguales que marcan los puntos de partida y las trayectorias educativas de los educandos.

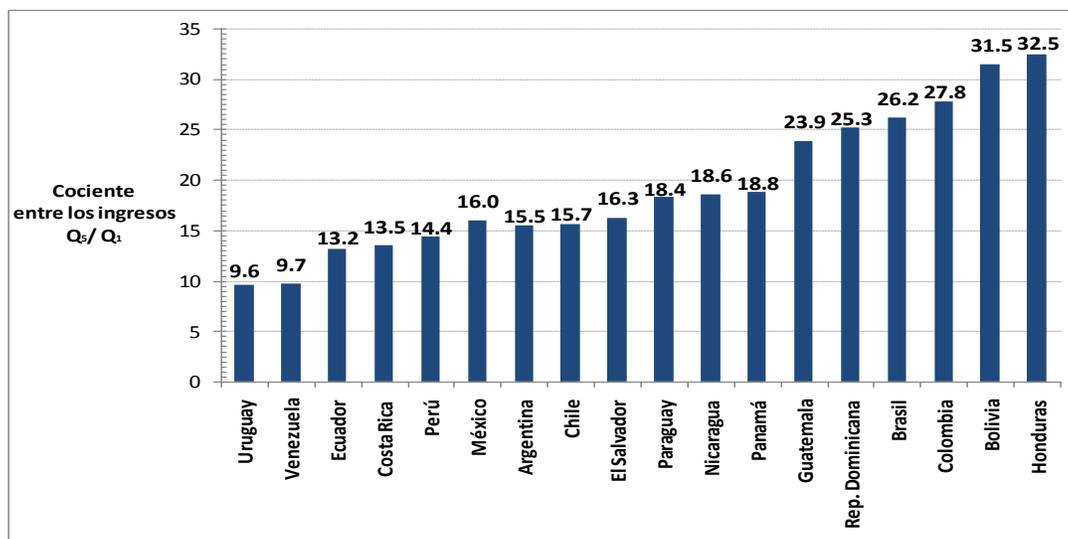
En este sentido, la bibliografía especializada ha señalado reiteradamente que la educación está condicionada tanto por el contexto socioeconómico como por el contexto político de la sociedad analizada, y que con frecuencia la educación reproduce la situación predominante, pero igualmente se ha señalado que la educación es un factor potencial de cambio social. Y por ello puede llegar a desempeñar un papel relevante en el objetivo de reducir las desigualdades sociales. Desde esta perspectiva, la escuela

puede y debe ofertar igualdad de oportunidades educativas, para evitar operar como un factor más de desigualdad social que acentúe la inequidad preexistente, sino más bien, operar como factor de transformación social al lograr que todos los grupos y estratos sociales tengan igualdad de oportunidades educativas que permitan, o al menos posibilitem, la reducción de las profundas brechas sociales.

1.1. Brechas sociales que marcan profundamente el contexto en que se desempeña el sistema educativo hondureño

Honduras presenta una de las mayores desigualdades sociales a nivel regional y mundial. Ya en el año 2004, el PNUD en su Informe sobre Desarrollo Humano a nivel mundial ubicaba a Honduras en el lugar 105° de 121 posibles en términos de desigualdad social, en peor situación que sus países vecinos El Salvador (94ª), México (104ª), Costa Rica (98ª) y Guatemala (100ª). De forma análoga, la CEPAL en su "Panorama Social de América Latina 2009", ubica a Honduras en el último lugar de la región al comparar los ingresos entre el quintil más rico y el más pobre (Pág. 89). Es en este contexto de extrema desigualdad social que se analizan los problemas de equidad y calidad del sistema educativo hondureño. Considerando en el análisis, no sólo el acceso al sistema, sino los insumos que se proveen y los resultados de aprendizaje alcanzados. Para la definición de los grupos de educandos a comparar se partió de los principios establecidos en la Convención sobre Derechos del Niño, aprobada en 1989 a iniciativa de las Naciones Unidas. En el preámbulo de la Convención se dice que "toda persona tiene todos los derechos y libertades ya enunciados, sin distinción alguna, por motivos de raza, color, sexo, idioma, religión, opinión política o de otra índole, de origen nacional o social, posición económica, nacimiento o cualquier otra condición".

GRÁFICO 1. AMÉRICA LATINA (18 PAÍSES): RELACIÓN DE INGRESOS ENTRE EL QUINTIL MÁS RICO Y EL QUINTIL MÁS POBRE. 2006-2008



Fuente: Elaboración propia en base a Comisión Económica para América Latina y el Caribe (CEPAL), sobre la base de tabulaciones especiales de las encuestas de hogares de los respectivos países. Informe: **Panorama Social de América Latina 2009**. Capítulo I. Anexo. p. 88-89.

2. EQUIDAD EN EDUCACIÓN Y DESIGUALDADES SOCIOECONÓMICAS

Si bien se reconoce el papel de la escuela para promover la equidad, dado que las personas que no acceden al sistema educativo también tienen problemas para acceder al mercado laboral; los estudios de los últimos veinte años vienen destacando que quienes no tienen una educación de calidad, entendida como la adquisición de un conjunto de competencias deseables, no pueden alcanzar la “plena ciudadanía”, al estar limitados para ejercer y defender sus derechos y tener acceso a la participación en los bienes sociales y culturales. Lo novedoso es el hecho de relevar la calidad sobre el acceso como condición necesaria: “La condición de ciudadano comprende el “currículum básico” indispensable que todos los ciudadanos han de poseer al término de la escolaridad obligatoria. Esto se define como un capital cultural mínimo y activo de competencias necesarias para moverse e integrarse en la vida colectiva; es decir, aquel conjunto de saberes y competencias que posibilitan la participación activa en la vida pública, sin verse excluido o con una ciudadanía negada” (Bolívar A., 2005).

En relación con el primer aspecto señalado, el acceso a la educación formal, **los datos nacionales indican claramente cómo las profundas desigualdades sociales se reflejan en las oportunidades educativas**. La condición de analfabeta es propia de las personas de los quintiles más pobres: Quintil 1 = 37% de analfabetismo; Quintil 2 = 27% de analfabetismo, etc., mientras que es una situación más bien “rara” en los grupos de mayores ingresos, Quintil 5 = 6% de analfabetismo. El análisis detallado por nivel de escolaridad, como se presenta en la siguiente tabla, muestra la relación directa entre quintil de ingresos del grupo de población y el acceso a los diferentes niveles del sistema educativo. De manera que del grupo de población adulta con escolaridad de nivel superior en Honduras (7.3% según EPH 2008), sólo el 1.5% corresponde a personas del quintil 1 de menores ingresos, mientras que el 77.6% corresponde al quintil 5 de mayores ingresos.

TABLA 1. DISTRIBUCIÓN DE LA POBLACIÓN ADULTA POR QUINTIL DE INGRESO Y NIVEL DE ESCOLARIDAD ALCANZADO. 2008

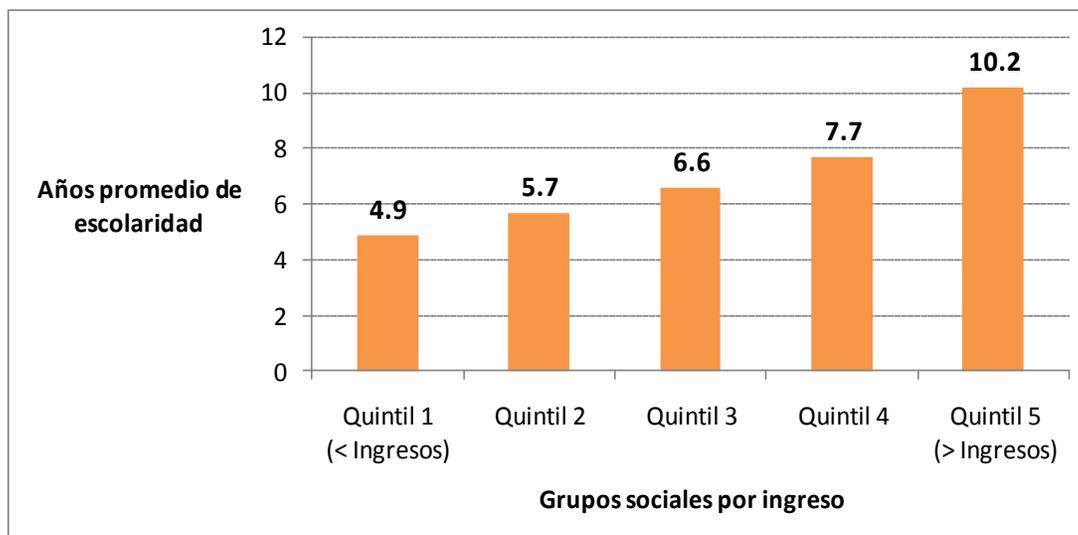
Nivel Educativo	Quintil 1 (20% de < ingreso)	Quintil 2	Quintil 3	Quintil 4	Quintil 5 (20% de > ingreso)
Ninguno	36.6%	27.0%	18.45	12.2%	5.7%
I y II Ciclo Básico	21.1%	22.7%	22.7%	20.6%	12.9%
III Ciclo Básico	5.2%	13.5%	23.0%	31.2%	27.1%
N. Medio	2.6%	7.3%	15.9%	28.5%	45.8%
Superior	1.5%	1.6%	4.6%	14.7%	77.6%

Fuente: INE (2008) **Encuesta Permanente de Hogares 2008**. Honduras.

Los años de escolaridad promedio de la población adulta también están estrechamente asociados con el nivel de ingresos de los sujetos, de manera que, en promedio, un adulto hondureño del quintil de mayores ingresos (Q_5) tiene más del doble de años de escolaridad que los hondureños del quintil de menores ingresos (Q_1), como se muestra en el siguiente gráfico. Esta desigualdad extrema entre dos grupos de población hondureños en años promedio de escolaridad (más de 2.08 veces), es mayor que la diferencia entre la media de países de Europa respecto a la media de países de Latinoamérica.

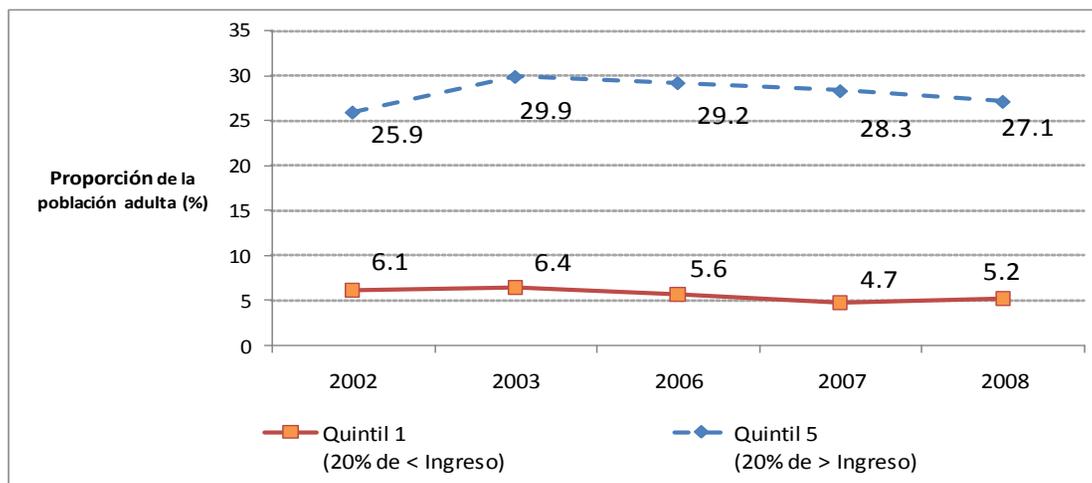
Y esta situación de amplia inequidad en el acceso a la educación formal no parece ir mejorando significativamente durante la última década, pese a los avances en matrícula neta que se vienen alcanzando en los últimos años para la educación básica. Como se aprecia en el siguiente gráfico, la composición del grupo de población adulta que alcanza el Tercer Ciclo de Educación Básica entre los años 2002-2008 apenas ha mejorado en el quintil 5 (mayores ingresos), pero incluso ha disminuido en el caso del quintil 1 (menores ingresos)

GRÁFICO 2. AÑOS PROMEDIO DE ESCOLARIDAD DE LA POBLACIÓN ADULTA SEGÚN SU QUINTIL DE INGRESOS. 2008



Fuente: INE (2008) Encuesta Permanente de Hogares 2008. Honduras.

GRÁFICO 3. EVOLUCIÓN DE LA PROPORCIÓN DE LA POBLACIÓN ADULTA CON ESCOLARIDAD DE III CICLO DE EDUCACIÓN BÁSICA EN EL PERÍODO 2002-2008



Fuente: Elaboración propia en base a INE (2008). Encuesta Permanente de Hogares 2002, 2003, 2006, 2007 y 2008.

Además de las profundas desigualdades en el acceso a la educación formal, esta situación se ve agravada por las desigualdades en los resultados de aprendizaje que claramente aparecen asociados a las condiciones socioeconómicas de los estudiantes. La siguiente tabla muestra cómo a menor condición socioeconómica de los alumnos, menores son los resultados de aprendizaje que se muestran en la evaluación externa. Estos resultados son ratificados al comparar los niveles de aprendizajes alcanzados por los alumnos de los centros de administración privada con respecto a los públicos. En relación con esta última comparación, además de las desigualdades en los resultados, se presentan diferencias importantes en los procesos de aula, en particular en relación con la variable tiempo de aprendizaje. Aunque no se

dispone de promedios nacionales, datos parciales indican que tanto en la duración de la jornada (de 6 a 7.5 horas diarias en los centros privados de "élite" respecto a 4.5 - 5 horas en las escuelas públicas), como en la cantidad de días de clase del calendario escolar (de cerca de 190 a 200 días de clase al año en los centros privados de "élite" respecto a 140 - 160 días en los centros públicos), los procesos de aula desarrollados en los centros privados aventajan significativamente a los públicos.

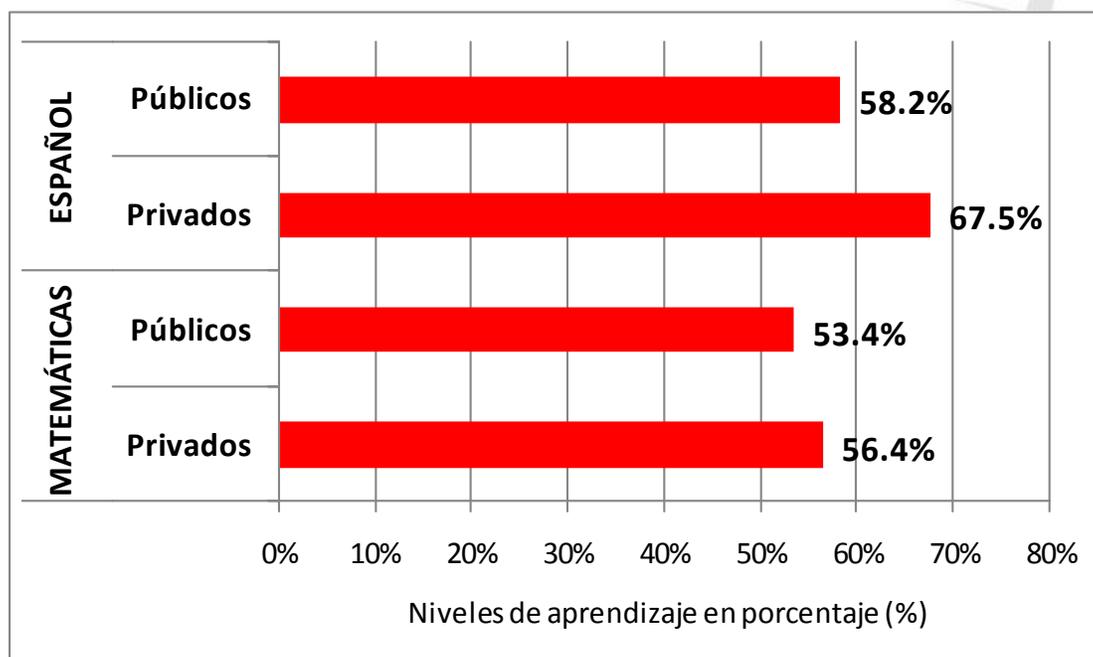
TABLA 2. PROPORCIÓN PROMEDIO DE RESPUESTAS CORRECTAS SEGÚN CONDICIÓN SOCIOECONÓMICA DE ALUMNOS DE SEXTO GRADO – 2004

Asignatura considerada	Condición socioeconómica de alumno		
	Desfavorable	Medio	Favorable
Español	38.8%	42.7%	49.7%
Matemáticas	35.1%	38.6%	42.6%
C. Naturales	39.1%	44.1%	51.1%

Fuente: SE-UMCE (2005) **Informe Comparativo Nacional del Rendimiento Académico 2002-2004**. Honduras

En el mismo plano de análisis, diferencias por nivel socioeconómico, los resultados de evaluación externa de los aprendizajes del 2008, confirman la tendencia marcada en los procesos anteriores (2002, 2004 y 2007), en el sentido de mejores resultados de parte de los alumnos de instituciones privadas en I y II Ciclo de Educación Básica. En el 2008, los niveles de aprendizaje mostrados por los alumnos de centros privados superaron 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 sólo dos o tres puntos porcentuales en promedio de los seis grados.

GRÁFICO 4. NIVELES DE APRENDIZAJE PROMEDIO DE LOS SEIS GRADOS, EN EVALUACIÓN EXTERNA 2008, SEGÚN TIPO DE ADMINISTRACIÓN DEL CENTRO EDUCATIVO

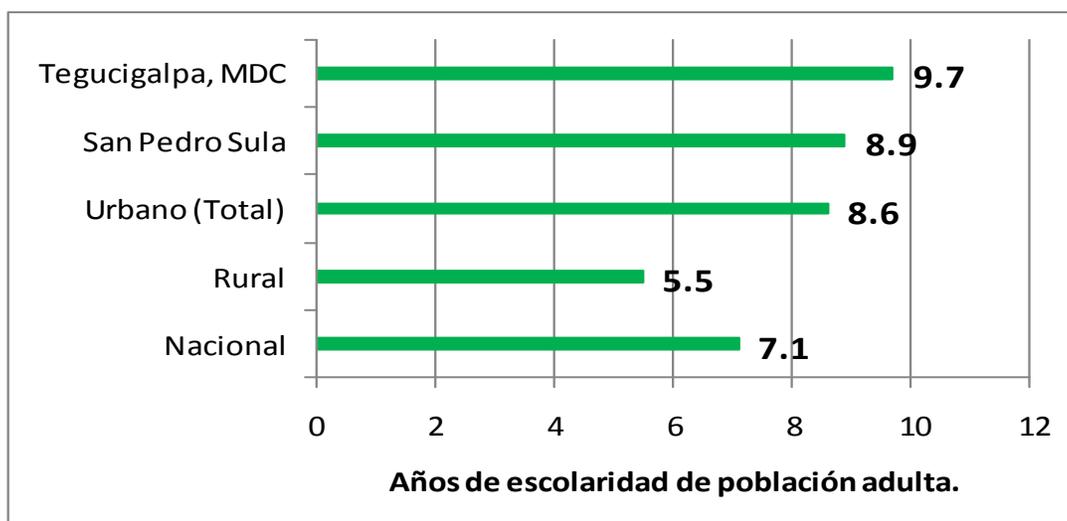


Fuente: SE-MIDEH (2009) **Informe Nacional de Evaluación de los Aprendizajes 2008**. Honduras.

3. EQUIDAD EN EDUCACIÓN Y DESIGUALDADES REGIONALES URBANO-RURALES

Una característica del Sistema Educativo hondureño a través de su historia ha sido presentar amplias desigualdades regionales urbano-rurales tanto en relación a cobertura como respecto a la cantidad y calidad de insumos proporcionados a los centros educativos, ello pese a que Honduras ha tenido una mayoría de población rural hasta el último censo del año 2000 (54% para ese año, 48.9% actualmente). Los problemas de acceso a la educación formal por parte de la población rural hondureña se reflejan en los niveles de escolaridad promedio de la población rural respecto a la urbana, evidenciándose además cómo los dos grandes centros urbanos del país, Tegucigalpa MDC y San Pedro Sula, alcanzan niveles educativos marcadamente superiores al resto de la población nacional, tal como se muestra en el gráfico 3.

GRÁFICO 5. AÑOS DE ESCOLARIDAD DE POBLACIÓN ADULTA SEGÚN REGIÓN DE RESIDENCIA. 2008



Fuente: Elaboración propia en base a INE (2008) **Encuesta Permanente de Hogares 2008**. Honduras.

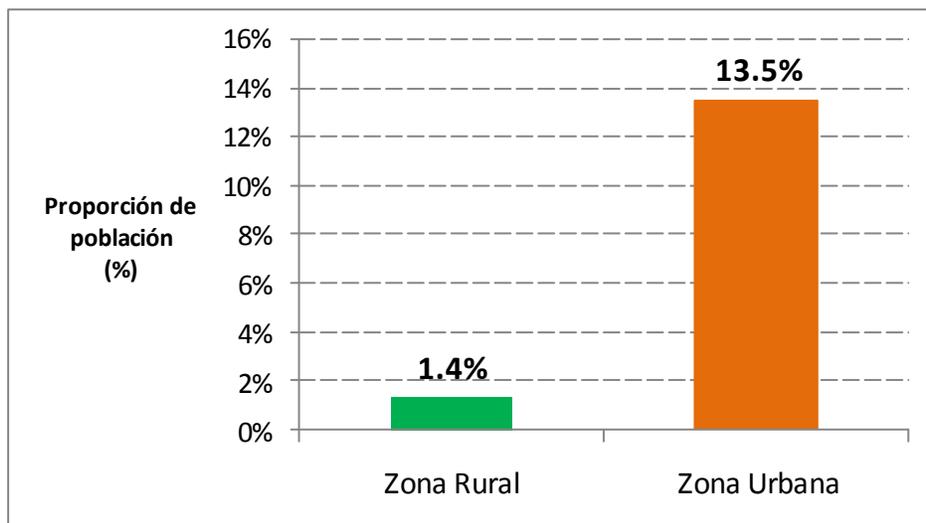
Las desigualdades de oportunidades en el acceso a la educación formal también se reflejan en la proporción de población adulta de cada región, urbana y rural, que alcanzan a culminar una carrera de nivel superior, siendo 7.3% la media nacional, pero poco más de 1% para la población del área rural, tal como se muestra en el siguiente gráfico.

Al mismo tiempo, las desigualdades en el acceso a la educación se reflejan en los años de escolaridad de la población económicamente activa PEA de cada región, en los que se evidencia que la población urbana alcanza un promedio de 8.9 años de educación respecto a 5.3 años para la población rural (INE, 2008).

Es relevante que las desigualdades en educación entre regiones urbanas y rurales también se expresan en los resultados de aprendizaje que los alumnos alcanzan en cada región. Los valores alcanzados en la evaluación externa de I y II Ciclo de Educación Básica a finales del 2008 ratifican esta tendencia. En este sentido, 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. Estos resultados son coherentes con los mostrados en las cuatro experiencias de evaluación externa realizados

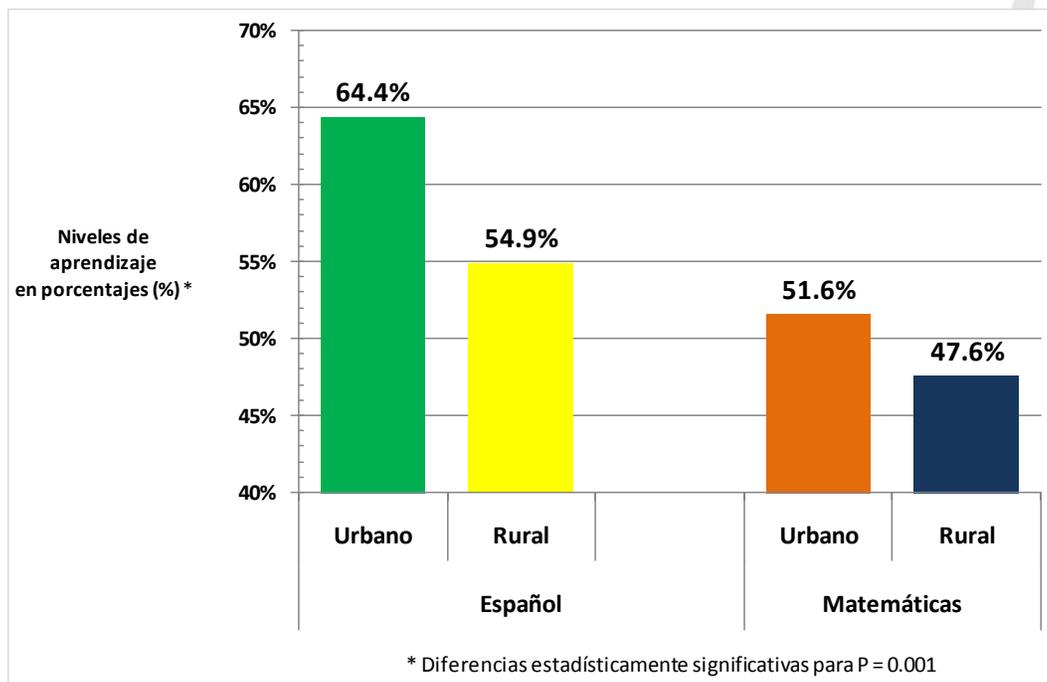
anteriormente (1997, 2002, 2004 y 2007), lo que puede interpretarse como la persistencia de una marcada inequidad en la calidad de los procesos educativos entre ambos grupos.

GRÁFICO 6. PORCENTAJE DE LA POBLACIÓN ADULTA CON NIVEL DE EDUCACIÓN SUPERIOR O UNIVERSITARIA, SEGÚN REGIÓN DE RESIDENCIA. 2008



Fuente: Elaboración propia en base a INE (2008) Encuesta Permanente de Hogares 2008. Honduras.

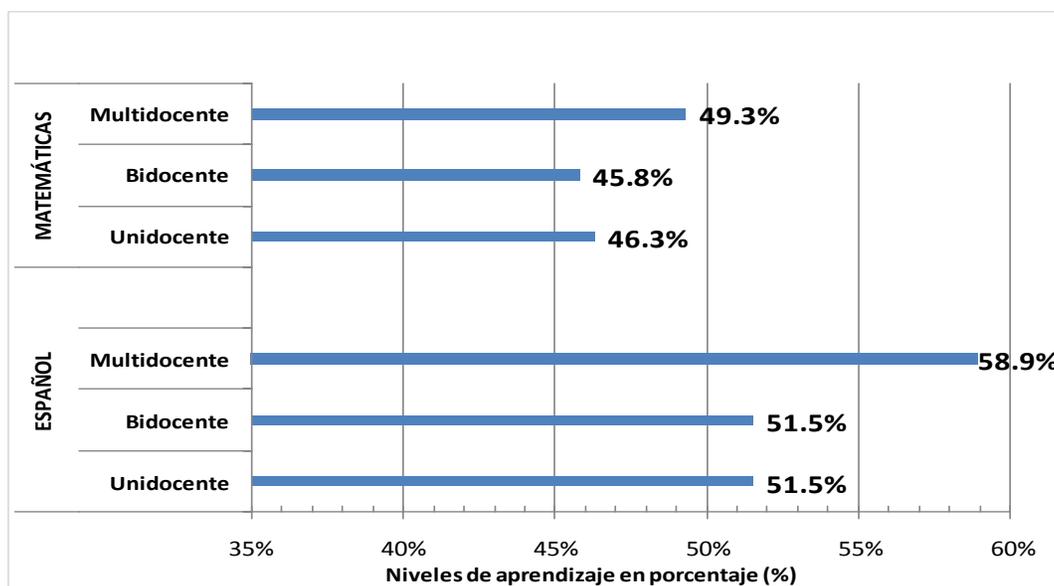
GRÁFICO 7. NIVELES DE APRENDIZAJE PROMEDIO PARA LOS SEIS GRADOS, SEGÚN ASIGNATURA Y REGIÓN DE UBICACIÓN DEL CENTRO EDUCATIVO. 2008



Fuente: SE-MIDEH (2009) Informe Nacional de Evaluación de los Aprendizajes 2008. Honduras. p. 19.

De forma análoga pueden interpretarse los niveles de aprendizaje alcanzados por los estudiantes en relación con el tipo de centro educativo al que asisten, según si cuentan con aulas multigrado o no. En este sentido, 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 los procesos educativos de los alumnos del área urbana respecto a los del área rural.

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



Fuente: Elaboración propia en base a INE (2008) Encuesta Permanente de Hogares 2008. Honduras.

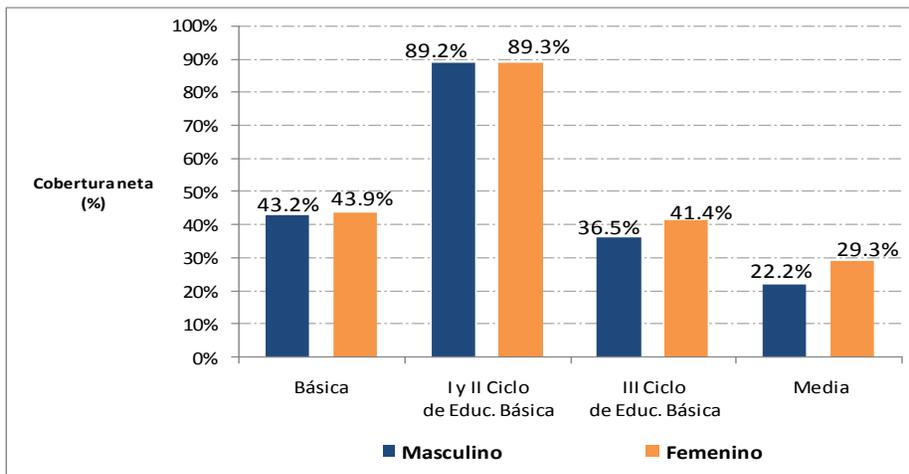
4. EQUIDAD EN EDUCACIÓN Y DESIGUALDADES DE GÉNERO

En términos generales, los sistemas educativos en América Latina no presentan **desigualdades** en el acceso en relación con el género, pero sí en la permanencia y finalización de estudios que en algunos países afecta a los varones por sobre las mujeres. No obstante, estudios regionales señalan que existen diferencias de oportunidades laborales entre varones y mujeres, que no están relacionadas con el nivel socioeconómico de los hogares, sino con el nivel educativo de los padres. En Latinoamérica tanto en las zonas urbanas como en las rurales las mujeres acumulan una ligera ventaja años de estudio respecto a varones. Sin embargo, entre quienes abandonan la educación, los datos indican que cerca de la mitad de las mujeres pasan a realizar labores domésticas en el hogar, mientras que casi la totalidad de los varones logra ingresar al mercado de trabajo, lo que les otorga más oportunidades de bienestar (CEPAL, 1998).

Afortunadamente, en Honduras la desigualdad de género en educación no es un problema importante en tanto los datos de la UPEG de la Secretaría de Educación plantean una relativa igualdad de proporción de matrícula, con pequeña ventaja para las niñas en los primeros ciclos (51.0%), ventaja que se incrementa

en el tercer ciclo y que además, es una constante en todos los departamentos del país para este ciclo y para los subsiguientes (media y superior).

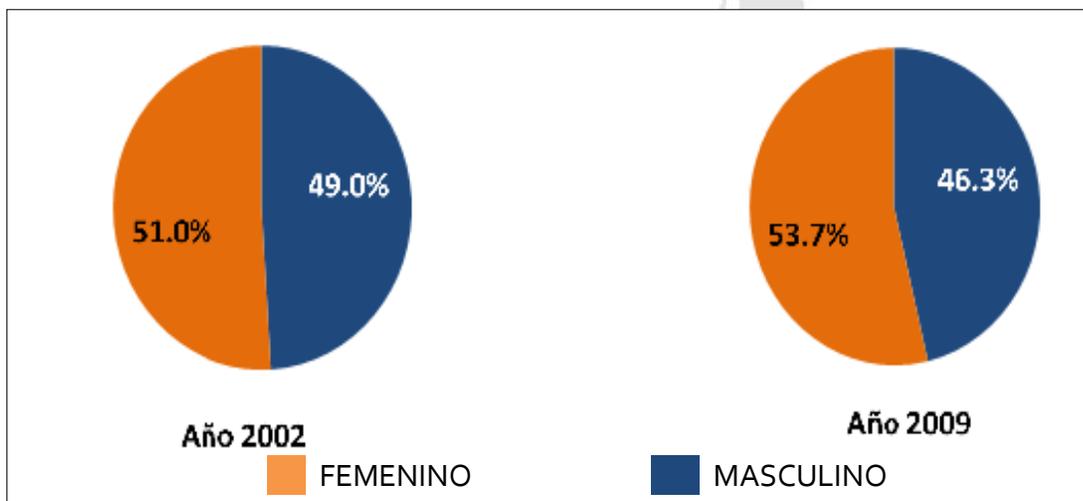
GRÁFICO 9. COBERTURA NETA POR NIVELES EDUCATIVOS Y EL GÉNERO DE LOS ESTUDIANTES. 2008



Fuente: INE (2008) **Encuesta Permanente de Hogares 2008**. Honduras.

En el caso del nivel superior de educación, el proceso de incremento y rebasamiento de la proporción de la matrícula femenina sobre la masculina ha sido más lento. En el año 2002 la diferencia era ya favorable al género femenino pero de una magnitud de sólo dos puntos porcentuales (51.0% respecto a 49.0%), en el 2006 la diferencia se incrementa a seis puntos porcentuales (53.0% respecto a 47%), y para el 2008 la distancia porcentual se ha elevado hasta 7.4 puntos (53.7% respecto a 46.3%). Esta última diferencia es mayor que cualquiera de las correspondientes a los niveles de Pre-Básica, Básica y Media en la actualidad, hecho que puede marcar tendencias en el futuro cercano para el mercado laboral de nivel universitario.

GRÁFICO 10. PROPORCIÓN DE MATRÍCULA DE NIVEL SUPERIOR, SEGÚN GÉNERO DEL ALUMNADO. 2002-2008



Fuente: Elaboración propia en base a INE (2008). **Encuesta permanente de Hogares**, años 2002 y 2008.

El hecho de que actualmente la matrícula femenina supere a la masculina en todos los niveles educativos del sistema nacional, se refleja ya en las proporciones de la población en condición de analfabetismo y en los años de escolaridad por género. Es relevante destacar que incluso entre la población de la región rural, la escolaridad de la mujer hondureña es ligeramente superior a la del hombre, como se aprecia en la siguiente tabla.

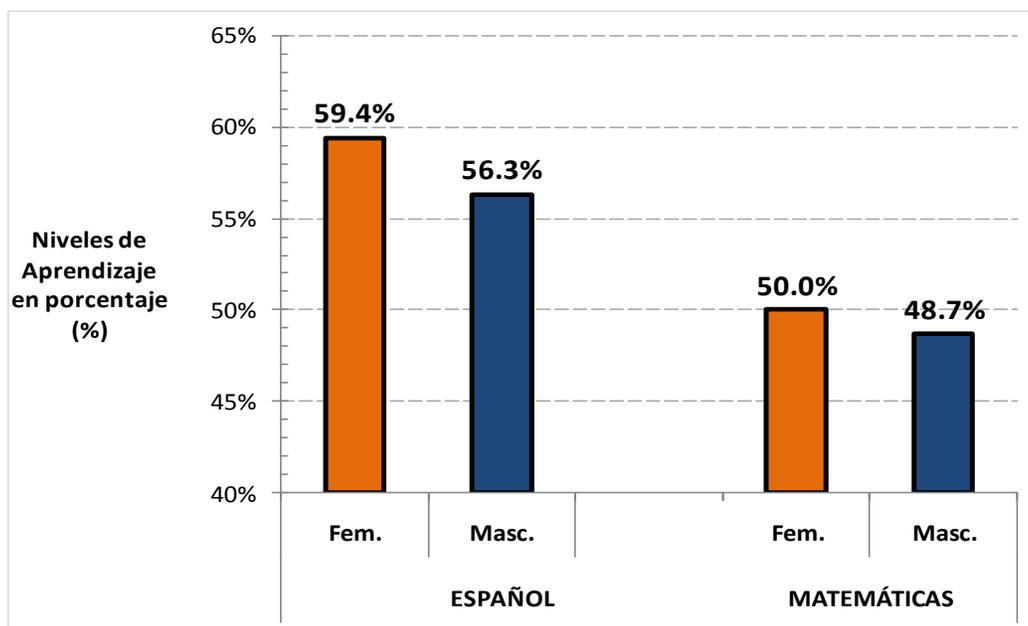
TABLA 3. AÑOS DE ESCOLARIDAD DE LA POBLACIÓN ADULTA POR GÉNERO Y REGIÓN. 2008

Categoría	Masculino	Femenino
Urbano	8.5%	8.6%
Rural	5.4%	5.7%
Nacional	7.1%	7.3%

Fuente: INE (2008) Encuesta Permanente de Hogares 2008. Honduras.

En los resultados de aprendizaje también se expresa la pequeña ventaja de las alumnas respecto a los alumnos. En la evaluación externa del 2008 para I y II Ciclo de Educación Básica, 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 resultados de aprendizaje.

GRÁFICO 11. NIVELES DE APRENDIZAJE PROMEDIO DE LOS SEIS GRADO SEGÚN ASIGNATURA Y GÉNERO DEL ALUMNADO



Fuente: SE-MIDEH (2009) Informe Nacional de Evaluación de los Aprendizajes 2008. Honduras. p. 21.

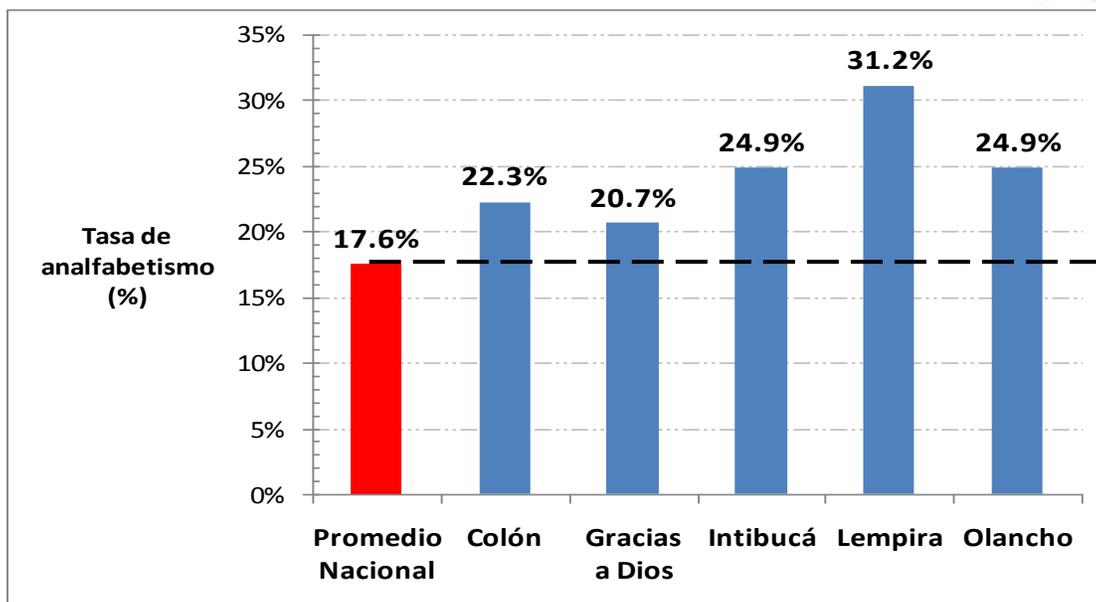
5. EQUIDAD EN EDUCACIÓN Y GRUPOS ÉTNICOS MINORITARIOS

Una de las formas de inequidad en educación que frecuentemente se analiza en los estudios internacionales es la referida a los grupos étnicos minoritarios. Se estima que en América Latina y el Caribe hay entre 40 y 50 millones de pobladores indígenas, con una variedad de entre 400 y 600 idiomas amerindios diferentes. Un buen número de pobladores indígenas vive en la zona rural o en zonas aisladas. Generalmente, las mayores tasas de analfabetismo y analfabetismo funcional se dan entre las mujeres indígenas.

Honduras presenta una rica diversidad étnica y cultural, dispersa por las diferentes partes del territorio nacional. Numerosas fuentes apuntan a que la composición de la población está distribuida en aproximadamente un 10% de población afrodescendiente e indígena y un 90% de mestizos. Estos grupos étnicos se distribuyen en 15 de los 18 departamentos del país, habitando 349 comunidades, siendo Gracias a Dios, Colón, Olancho, Lempira e Intibucá los cinco departamentos con mayor concentración indígena y afrodescendiente (Amaya J.A., 2009). Y dada la carencia de información pormenorizada a nivel de municipio y/ o aldea para estos grupos étnicos, en el presente análisis se utilizan los datos educativos referidos a estos cinco departamentos para, en términos gruesos, comparar sus indicadores educativos.

En relación con el acceso a la educación formal, es claro que los cinco departamentos que albergan la mayor concentración de población indígena y afrodescendiente del país presentan índices de analfabetismo superiores a la media nacional (17.6%), todos con tres o más puntos porcentuales arriba de ese valor, incluyendo además al departamento más afectado por este flagelo sociocultural, Lempira, con casi un tercio de su población en situación de analfabetismo.

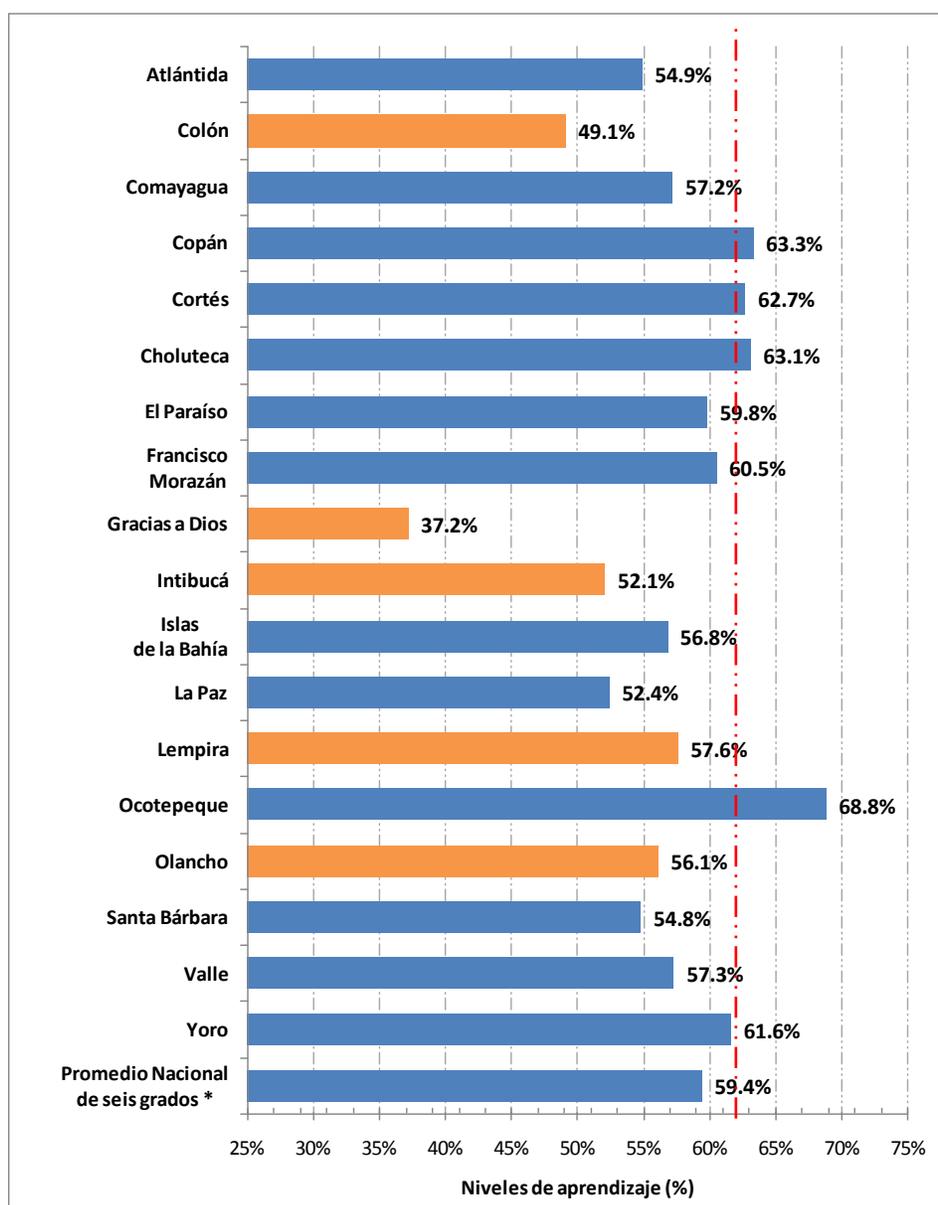
GRÁFICO 12. TASAS DE ANALFABETISMO POR DEPARTAMENTO, COMPARANDO LOS QUE ALBERGAN LA MAYOR PROPORCIÓN DE POBLACIÓN INDÍGENA Y AFRODESCENDIENTE



Fuente: Elaboración propia en base a PNUD (2009). **Informe sobre Desarrollo Humano Honduras 2008/ 2009**. Honduras, p. 411-416.

Los resultados de aprendizaje también muestran la situación desventajosa en que se encuentran los departamentos que albergan la mayor proporción de población indígena y afrodescendiente. De acuerdo con los reportes de la Secretaría de Educación para la evaluación externa del 2008, los resultados promedio de los cinco departamentos referidos están entre los más bajos, incluyendo los últimos tres lugares (16°, 17° y 18°) tanto para las pruebas de Matemáticas como de Español en los seis grados de I y II Ciclo de Educación Básica.

GRÁFICO 13. RESULTADOS PROMEDIO DE APRENDIZAJE EN ESPAÑOL POR DEPARTAMENTO, COMPARANDO LOS CINCO DEPARTAMENTOS QUE ALBERGAN LA MAYOR PROPORCIÓN DE POBLACIÓN INDÍGENA Y AFRODESCENDIENTE. EVALUACIÓN EXTERNA 2008



Fuente: SE-MIDEH (2009). **Evaluación Nacional de los Aprendizajes 2008**. Honduras, p. 17.

6. HALLAZGOS Y RETOS MÁS IMPORTANTES

La panorámica aquí presentada respecto a la educación hondureña permite identificar algunos resultados especialmente significativos, así como algunos de los desafíos más importantes que se presentan en relación con la temática de Equidad en la Educación Hondureña.

6.1. El sistema educativo reproduce las profundas desigualdades sociales de la sociedad hondureña

Actualmente tanto los indicadores educativos de acceso (por ejemplo, escolaridad promedio por quintil de ingreso), como los de resultados (promedios de aprendizaje en evaluación externa según nivel socioeconómico del alumno), se corresponden claramente con las profundas desigualdades sociales que padece la sociedad hondureña (la mayor en América Latina según CEPAL 2009 si se aplica el indicador de cociente de quintiles Q_5/Q_1). Los datos indican que el sistema educativo está reflejando, y posiblemente reproduciendo, las amplias brechas sociales imperantes. Ninguno de los indicadores analizados altera la correspondencia desigualdad social- desigualdades educativas. Esta crítica situación plantea el desafío urgente de diseñar e implementar políticas de reducción de las desigualdades en educación, focalizadas en los grupos de menores ingresos.

6.2. Las desigualdades en educación entre regiones urbano-rurales y respecto a los grupos étnicos minoritarios, son acentuadas

La desigualdad en educación entre zonas urbanas y rurales es particularmente relevante en el caso de países que, como Honduras, casi la mitad de la población es rural y además en esta región confluyen otros factores de vulnerabilidad tales como la pobreza y limitadas condiciones de acceso a servicios de salud. Al respecto, tanto los indicadores educativos de acceso tales como cobertura neta por nivel, como los de resultados educativos (como promedios de aprendizaje en evaluación externa), evidencian acentuadas desventajas para los alumnos del área rural y/o grupos indígenas o afrodescendientes (que también habitan predominantemente en áreas rurales). Las desigualdades en el acceso a la educación son mayores a medida que se asciende en los niveles educativos, siendo los jóvenes de 13 años en adelante los que más dificultades enfrentan para continuar sus estudios en secundaria o en el nivel superior.. Es prioritario el desarrollo de políticas compensatorias focalizadas para estos grupos de población, de las cuales el Programa de Educación Bilingüe Intercultural EBI es un ejemplo a seguir.

6.3. El Sistema Educativo Hondureño no presenta desigualdades de género que afecten al alumnado femenino

Los datos de la última década indican que en el sistema educativo hondureño se han superado las desventajas de acceso y resultados de aprendizaje vinculados al género femenino. Todos los indicadores de acceso a los diferentes niveles del sistema muestran una creciente ventaja de las alumnas respecto a los alumnos (que en educación media llega a ser de 7 puntos porcentuales), mientras que los resultados de aprendizaje alcanzados en la evaluación externa también se muestra una ligera superioridad de las estudiantes respecto a sus pares masculinos, tanto en Matemáticas como en Español. En esta área en particular, el sistema educativo hondureño ha alcanzado la meta señalada por el Foro Mundial sobre la Educación en Dakar, sobre "Suprimir las disparidades entre los géneros en la enseñanza primaria y

secundaria en particular garantizando a las niñas un acceso equitativo a una educación básica de buena calidad, así como un buen rendimiento”.

6.4. Deben reorientarse las políticas que buscan la equidad en el sistema educativo hondureño

Durante las dos últimas décadas se han desarrollado en el país numerosas políticas en educación orientadas a favorecer a los sectores de menores recursos y población rural o de grupos étnicos minoritarios. Se pueden contabilizar un número considerable de instituciones y proyectos públicos y privados que han venido trabajando el tema educativo en poblaciones vulnerables. Sin embargo, la información disponible tanto respecto al acceso como a los resultados educativos indica que la equidad educativa es aún una meta lejana para la sociedad hondureña. Debe considerarse en la toma de decisiones que tanto los datos disponibles en Honduras como la experiencia internacional respecto a las políticas orientadas a la equidad en educación indican claramente que el impacto es mayor si las intervenciones se focalizan en los grupos de población más desfavorecidos. “Es necesario iniciar verdaderas políticas de acción afirmativa, de discriminación positiva, que busquen efectivamente igualar o dar más recursos y atención a los estudiantes de los grupos de menores ingresos” (Reimers F., 2000: 44). En el actual contexto hondureño de profundas y crecientes desigualdades sociales amén de una crónica escasez de recursos, es urgente tomar medidas al respecto en las políticas educativas orientadas a la equidad.

REFERENCIAS BIBLIOGRÁFICAS

- Amaya, J.A. (2009). *La participación de los movimientos sociales indígenas y afrohondureños en el FNR*. Honduras.
- Blanca, R. y Cusato, S. (2004). *Desigualdades educativas en América Latina: Todos somos responsables. Escuelas de Calidad en condiciones de pobreza*. Chile: Universidad Alberto Hurtado – BID.
- Bolívar, A. (2005). Equidad educativa y teorías de la justicia. *Revista Electrónica sobre Calidad, Eficacia y Cambio en Educación*, 3(2).
- Calvo, G. (2006). *La preocupación por la equidad y la formación docente: A propósito de la Eficacia Escolar*. Colombia.
- Carnoy, M. (2005). La búsqueda de la igualdad a través de las Políticas Educativas: Alcances y Límites. *Revista Electrónica Iberoamericana sobre Calidad y Cambio Educativo*, 3(2).
- CEPAL (1998). *Panorama social de América Latina 1998*. Chile.
- CEPAL (2009) *Panorama social de América Latina 2009*. Chile.
- Cohen, E. (2002). Educación, Eficiencia y Equidad: Una difícil convivencia. *Revista Iberoamericana de Educación*, 30, pp. 105-124.
- ESA Consultores (2002). *Análisis alternativo proyección de costos para la reducción de la pobreza en el tercer ciclo de educación básica en Honduras*. Tegucigalpa: USAID.
- Franco, R. (1995). Viabilidad económica e institucional de la reforma educativa en Chile. *Serie Políticas Sociales*, 11. Santiago de Chile: CEPAL.
- Foro Mundial de la Educación (2000). *Marco de trabajo de Dakar*.

- INE (2008). *Encuesta Permanente de Hogares 2008*. Honduras.
- Marchesi, A. (2000). Un sistema de indicadores de desigualdad educativa. *Revista Iberoamericana de Educación*. 23(monográfico Equidad en Educación).
- Marshall, J., Mejía, M., Aguilar, C. (2005). *Calidad y Eficacia en una escuela alternativa: Lecciones de la experiencia de EDUCATODOS en Honduras*. Tegucigalpa: USAID.
- PREAL (2006). *Educación y brechas de equidad en América Latina*. Chile: Fondo de Investigaciones Educativas.
- PNUD (2004). *Informe Mundial sobre Desarrollo Humano*.
- PNUD (2006). *Informe Mundial sobre Desarrollo Humano 2006*. Honduras.
- PNUD (2009). *Informe sobre Desarrollo Humano Honduras 2008/ 2009*. Honduras.
- Reimers, F. (2000). Educación, desigualdades y opciones de política en América Latina en el siglo XXI. *Revista Iberoamericana de Educación*, 23(monográfico Equidad en Educación).
- Reimers, F. (2002). *Distintas escuelas, diferentes oportunidades*. Madrid: La Muralla.
- SE/ UPEG (2008). *Bases de datos de Educación Básica*. Tegucigalpa.
- SE/ PROHECO (2009). *10 años llevando riqueza educativa a comunidades con pobreza extrema*. Tegucigalpa.
- SE/ PROHECO (2008). *Memoria 2008. Programa Hondureño de Educación Comunitaria*. Tegucigalpa.
- Sen, A. (1995). *Nuevo examen de la desigualdad*. Madrid: Alianza.
- SE-UMCE (2005). *Informe Comparativo Nacional de Rendimiento Académico: 2002-2004*. Honduras.
- SE-MIDEH (2009). *Informe Nacional de Evaluación de los Aprendizajes 2008*. Honduras.
- UMCE (2000). Informe global del rendimiento académico de 2º y 3º grados. *Proyecto Hondureño de Educación Comunitaria*. Honduras.
- UMCE (2000b). Informe en Español de 2º y 3º grados. *Proyecto Hondureño de Educación Comunitaria (PROHECO)*. Honduras.
- UMCE/ UPNFM (2002). *Evaluación del séptimo grado de educación básica programa EDUCATODOS: Español y Matemáticas*. Tegucigalpa.
- Willie, C. (2002). Excelencia, equidad y diversidad en educación. Reimers, F. *Distintas escuelas, diferentes oportunidades: retos para la igualdad de oportunidades en Latinoamérica*. pp. 99-112. Madrid: La Muralla.

STAKE, ROBERT E. (2010). INVESTIGACIÓN CUALITATIVA: EL ESTUDIO DE CÓMO FUNCIONAN LAS COSAS. NEW YORK: THE GUILFORD PRESS

STAKE, ROBERT E. (2010). QUALITATIVE RESEARCH: STUDYING HOW THINGS WORK. NEW YORK: THE GUILFORD PRESS

Zulma Perassi

Revista Iberoamericana de Evaluación Educativa 2010 - Volumen 3, Número 3

<http://www.rinace.net/riee/numeros/vol3-num3/reseña.pdf>



Robert E. Stake es Director del Centro de Investigación Instruccional y Evaluación Curricular en la Universidad de Illinois en Urbana-Champaign. Es uno de los investigadores educacionales autores de la teoría y práctica para la evaluación de programas educativos en la década de 1960. Su enfoque de la evaluación -a menudo en forma de estudios de casos- enfatiza el estudio de la experiencia en el aula, la interacción personal, los procesos institucionales y los contextos.

Entre sus publicaciones más conocidas se encuentran:

Multiple Case Study Analysis, Standards-based Responsive Evaluation, Quieting Reform: Social Science and Social Action in an Urban Youth Program, Evaluating the Arts in Education: A Responsive Approach y The Art of Case Study Research.

La obra que aquí se presenta ha sido recientemente publicada y aún no fue traducida al español.

El propósito principal que posee el autor al escribir este libro es crear situaciones de aprendizaje, provocar en el docente o en el lector, mejores experiencias. Él intenta construir diferentes "cuñas" –cual sólidos postes donde se amarran los barcos en un puerto- ofreciéndole a cada uno la posibilidad de atar en ellos sus preguntas, ejemplos, explicaciones... constituyendo una atractiva invitación a participar del diálogo con el autor.

Stake aborda la investigación cualitativa procurando que cada lector se comprometa personalmente, que recupere las experiencias propias y aquellas preguntas formuladas a lo largo de su vida. Por cierto que para desarrollar proyectos de investigación es necesario ser más disciplinado y comprometido con los temas de estudio, aunque ello no significa necesariamente un proceder más formal.

Este libro se orienta a la comprensión del mundo social y profesional, gestándose a partir de la atención de aquello que la gente "hace y dice", contemplando cómo manifiestan su productividad y su amor. El texto no pretende definir cómo deberían funcionar las cosas, sino que intenta ayudar a mejorar la habilidad para analizar cómo están funcionando ellas. En esta línea, la preocupación está centrada en profundizar cómo algo particular está funcionando, más que en considerar la marcha de las cuestiones generales. Trabajar en generalizaciones amplias requiere extensos estudios y el apoyo de métodos cualitativos y cuantitativos y no es este el sentido del presente trabajo. Aquí se enfatiza la comprensión de lo que está sucediendo.

Cuando las complejidades no son comprendidas, los problemas suelen ser tratados superficialmente, buscando no demorar la actuación para que las cosas mejoren. Puede el investigador quejarse de los problemas mientras desarrolla su labor, pero tomar posición temprana para remediar situaciones particulares, generalmente aleja a la investigación de sus propósitos centrales.

La propuesta del libro acompaña al lector a trabajar a su manera, sin embargo exige mayor paciencia a medida que más experta sea la persona en el funcionamiento de las cosas. Se trata de construir nuevos caminos, de pensar en profundidad un estudio desde el principio al final, de comprender lo investigado.

En este texto, el autor moldea cada concepto a la luz de su experiencia profesional, de su historia laboral, del desplazamiento que el mismo vivió desde los métodos cuantitativos hacia la investigación cualitativa. Enfatiza los contextos de acción y procura desechar estereotipos.

En la introducción de este libro, Stake expone una secuencia de proyectos (desde el Proyecto A hasta el Proyecto T) que ayudan a comprender los conceptos y métodos desarrollados en los doce capítulos del

texto. Dichos proyectos operan como facilitadores del proceso de investigación cualitativa, ofreciendo algunas acciones concretas para trabajar en distintas fases del desarrollo.

Todo pensamiento científico - para este autor- es una mezcla de pensamiento cualitativo y cuantitativo. La investigación es en esencia una búsqueda de comprensión. En este sentido, se destacan como características claves del estudio cualitativo, el carácter interpretativo del mismo, su naturaleza empírica, experimental, el privilegio de la situacionalidad –cada lugar y tiempo delimitan contextos únicos-, su orientación personalista en tanto se ocupa de comprender las percepciones humanas, honrando la diversidad y emergiendo más de la gente, que de los propios investigadores.

Para Stake la investigación es *investigación interpretativa* y depende en gran medida de la habilidad que desarrollan los observadores para definir y redefinir el significado de lo que ven y oyen. Estas elaboraciones están estrechamente vinculadas a la adecuada comprensión de las condiciones del entorno, del contexto y de cada situación. Gran parte de la formación que requiere la investigación cualitativa es aprender a minimizar las fallas de las propias observaciones y afirmaciones que realiza el investigador. Por ello, la triangulación de datos, fuentes, perspectivas... permite incrementar la confianza en dichas interpretaciones.

El autor está interesado en la investigación de casos: en las especificidades, las narraciones, las situaciones y aquellos sucesos o episodios que develan la forma particular en que funcionan las cosas, es decir, en la investigación cualitativa denominada microanálisis.

Este texto es una invitación a descubrir "cómo funcionan las cosas" ya sea en las organizaciones, programas, comunidades o sistemas. Los temas que desarrolla Robert Stake a través de los doce capítulos del libro, recorren una secuencia que va desde la identificación de una pregunta de investigación, a la selección de métodos, atravesando por la revisión bibliográfica necesaria para acercarse al problema, la recopilación de datos y evidencias, el análisis e interpretación de los resultados, para concluir con la elaboración del informe final.

Esta producción, atravesada por un sentido ético e ilustrada con múltiples ejemplos de la nutrida vida profesional de su autor, convoca a profesores, estudiantes e investigadores a desplegar juntos con él, un interesante proceso investigativo.