

**LOS ALUMNOS BAJO LA LUPA:
LOS EXÁMENES EXTERNOS CON
CONSECUENCIAS INDIVIDUALES***

Bárbara Eyzaguirre

En Chile, el tema de la rendición de cuentas de los establecimientos educacionales, profesores y alumnos ha estado prácticamente ausente de la discusión pública. Sin embargo, la evidencia comparada muestra que examinar a los alumnos es una política común en la mayoría de los países de excelencia académica, tanto europeos como asiáticos, y recientemente se la ha comenzado a aplicar en los Estados Unidos. La teoría y la evidencia empírica apuntan a que los sistemas de rendición de cuentas de los alumnos afectan positivamente a la estructura de incentivos de estudiantes, padres y escuelas, y por ende al nivel de aprendizaje.

En este artículo se explora el marco general en el cual se inserta un sistema de rendición de cuentas de los estudiantes y se discuten las principales razones que justifican su implementación. También se analizan la evidencia empírica y los principales problemas que suscita la aplicación de este tipo de exámenes, incluyendo las soluciones ensayadas. Finalmente, se esboza una propuesta para introducir en nuestro país un sistema de rendición de cuentas con consecuencias individuales.

BÁRBARA EYZAGUIRRE. Psicóloga educacional especializada en desarrollo cognitivo, con experiencia en programas de mejoramiento de la calidad de la educación en sectores de escasos recursos. Fundadora y asesora pedagógica de la Escuela San Joaquín (Renca), perteneciente a la Fundación Marcelo Astoreca. Investigadora del Centro de Estudios Públicos.

* Versión revisada del trabajo publicado en *Documento de Trabajo* N° 331 (abril 2002) del Centro de Estudios Públicos. Se agradecen los valiosos aportes de Harald Beyer y Carmen Le Foulon, y también los comentarios de Arturo Fontaine T. y Salvador Valdés.

Estudios Públicos, 88 (primavera 2002).

Se han introducido en Chile, en las últimas décadas, una serie de medidas para mejorar la calidad de la educación. Sin embargo, hay una línea de innovaciones que no se ha abordado en forma decidida. Nos referimos a los mecanismos para exigir cuentas a los establecimientos, profesores y alumnos¹.

A partir de los años 80, la rendición de cuentas (*accountability*) y los temas de evaluación e incentivos han sido centrales en las discusiones sobre políticas educativas en los países desarrollados. La preocupación creciente por los resultados educativos obedece al “reconocimiento, ampliamente compartido, de que la educación es una inversión en capital humano que aporta ganancias reales a la sociedad y a la economía”². Dentro de estos aportes se incluye el impacto positivo de la educación en la distribución del ingreso. La evidencia indica que las diferencias de ingreso están determinadas en gran medida por la dispersión de conocimientos y destrezas³. Por estas y otras razones que no vienen al caso desarrollar aquí, los temas educacionales han adoptado un perfil político cada vez más alto, lo cual se ha traducido en un aumento del gasto público en el área y el consiguiente interés de exigir resultados a cambio.

La rendición de cuentas busca, en términos simples, exigir que los actores demuestren en qué medida están cumpliendo con los objetivos fijados. Para esto es necesario que primero se definan los objetivos, luego se establezcan las consecuencias que enfrentarán si cumplen o dejan de cumplir las metas, y en tercer lugar, se reúna la información que indique si los objetivos se están logrando o no.

En este artículo se abordará específicamente el tema de rendición de cuentas de los alumnos a través de exámenes externos con consecuencias individuales. Chile no cuenta con este tipo de exámenes, en cambio los países que alcanzan un nivel alto de aprendizaje sí los tienen. Implementar una medida de esta naturaleza puede contribuir a elevar los bajos estándares imperantes en el país.

En la primera sección se explora el marco general en el cual se inserta un sistema de rendición de cuentas de los estudiantes. Luego se

¹ Para una revisión de la reforma educacional en Chile, véase H. Beyer, B. Eyzaguirre y L. Fontaine, “Reseña de La Reforma Educacional Chilena” (2001), donde se resalta la ausencia de medidas de rendición de cuentas. El estudio en cuestión puede servir como referencia para comprender el lugar que ocupa la rendición de cuentas en el conjunto de intervenciones que se deben realizar para mejorar la calidad de la educación.

² Center for Education Research and Innovation, *Schools under Scrutiny* (1995), p. 15.

³ H. Glennester, “United Kingdom Education 1997-2001” (2001). Véase también, H. Beyer, “Educación y Desigualdad de Ingresos: Una Nueva Mirada” (2000).

define qué se entiende por ello y se discuten las principales razones que justifican su implementación. En la tercera parte se analiza la evidencia empírica y en la cuarta se presentan los principales problemas que suscita la aplicación de este tipo de exámenes, incluyendo las soluciones que se han experimentado. Finalmente se analiza en qué situación se encuentra este tema en Chile y se propone, para la discusión, un sistema de rendición de cuentas con consecuencias individuales.

I. LOS SISTEMAS DE RENDICIÓN DE CUENTAS EN EDUCACIÓN

A continuación se delinearán brevemente los principales parámetros que intervienen en un sistema de rendición de cuentas.

¿Quiénes tienen que rendir cuentas?

En un sistema complejo como el educacional son muchos los estamentos que participan y, según las investigaciones, todos ellos deben rendir cuentas si se aspira a un nivel de rendimiento alto⁴. En primer lugar, están los agentes del gobierno central, seguidos por los del gobierno local, luego los administradores de escuelas, los profesores y finalmente los alumnos.

¿A quién se le rinde cuentas?

La rendición de cuentas depende de cuán centralizado o descentralizado sea el sistema de educación. En algunos casos las escuelas dan cuenta directa al Estado, el que a su vez responde ante los votantes. En otros casos, las cuentas se dan directamente a los padres. En muchas naciones los establecimientos responden a más de una instancia, ya que prevalece la idea de que la educación es un bien público que afecta no sólo al interés individual del alumno y de la familia sino a la sociedad en su conjunto.

¿Cómo se definen los criterios o los estándares según los cuales se rinden cuentas?

En los países que tienen tradición de excelencia educacional pueden no estar formalizados los estándares, pero las personas encargadas de construir los instrumentos saben cuáles son los niveles de desempeño esperados.

⁴ Véase Helen F. Ladd (ed.), *Holding Schools Accountable: Performance-Based Reform in Education* (1996).

Hay un consenso implícito sobre los mínimos aceptables. Probablemente, los estándares altos son internalizados por cada una de las personas, durante su experiencia escolar, cuando tienen que superar cada una de las vallas académicas requeridas para obtener su certificación⁵. En países en que la calidad de la educación no es buena es necesario definir qué estándar será el requerido. En este punto es crucial que el currículo explicitamente los objetivos de logro y los respalde directamente en las pruebas de evaluación del rendimiento del sistema. Estados Unidos tiene una tradición de libertad curricular, sin embargo, ante la percepción de que su educación estaba en crisis aceleró el proceso de fijar estándares. Actualmente la *educación basada en estándares* permea al país. Sólo Iowa no tiene estándares oficiales, aunque se espera que sus distritos los establezcan. Marzano (1998, p. 1), un exponente reconocido del movimiento de desarrollo de habilidades de pensamiento, ha impulsado este cambio y afirma que “es un imperativo que los distritos y las escuelas planifiquen cuidadosamente sus estándares de manera que los efectos en las clases sean fruto del diseño y no de la espontaneidad”.

¿Cuáles son los instrumentos que se utilizan para verificar el cumplimiento de los estándares?

Entre las prácticas comunes encontramos las inspecciones, encuestas de indicadores y la utilización de pruebas de evaluación. La frecuencia de aplicación, las áreas cubiertas al evaluar y los segmentos de edad considerados difieren significativamente entre un país y otro.

¿Qué se hace con la información recopilada?

Hay países en los cuales las instancias centrales manejan la información de manera restringida para la toma de decisiones y hay otros que la difunden ampliamente. La tendencia actual es abrir esta información a los padres.

¿Cuáles son las medidas que se toman una vez que se conocen esos resultados?

En general los sistemas establecen un rango de desempeño que consideran adecuado. Ante un resultado satisfactorio, hay un reconocimiento

⁵ Los profesores tienden a exigir niveles similares a los que ellos fueron sometidos, a menos que se sientan muy aventajados respecto a los estudiantes que están educando, en cuyo caso pueden adoptar una actitud condescendiente y exigirles menos que lo que requerirían normalmente.

implícito de que están cumpliendo con la labor encomendada. En relación a los que se salen de los márgenes, los países que tienen buenos sistemas educacionales establecen una diversidad de medidas para incentivar la corrección de los déficits y para estimular a los que alcanzan niveles de excelencia. Entre las aplicadas a los alumnos encontramos: asistencia a clases de recuperación; aceleración, promoción o repitencia de cursos; entrega o retención del certificado de graduación; agrupación de los alumnos por rendimiento o por modalidades vocacionales (*tracking/streaming*); ingreso a instituciones de prestigio y asignación de becas para asistir a la universidad.

Para los profesores encontramos pago por mérito, premios de excelencia, asignación de mayores responsabilidades, o reasignación de tareas, obligación de reactualizar conocimientos volviendo a la universidad, inhabilitación para el cargo.

En el nivel de los establecimientos, la mera publicación de los puntajes obtenidos en evaluaciones estandarizadas se traduce en reconocimiento o en sanción social. Para potenciar estos efectos, algunos estados norteamericanos han recurrido a categorizar el desempeño de las escuelas en unos pocos conceptos que tienen más fuerza que los puntajes puros. Por ejemplo, en Texas las categorías son: “escuelas académicamente inaceptables, aceptables, eficientes y sobresalientes”. En Estados Unidos, 27 estados califican a sus escuelas sobre la base del puntaje en las pruebas y 16 estados incorporan, además, otro tipo de indicadores. Una práctica común en Inglaterra es difundir, además, el número de alumnos que una escuela logra posicionar en instituciones de enseñanza media y universitarias de prestigio. En los sistemas en que los padres pueden escoger los establecimientos, como en Chile, el prestigio adquirido debiera traer como consecuencia un aumento o disminución de matrículas, lo cual sería un incentivo importante para mejorar la calidad de la educación entregada. En Estados Unidos, donde los padres no pueden cambiar fácilmente a sus hijos de colegio, 14 estados han instituido medidas para evitar que los alumnos queden atrapados en forma indefinida en establecimientos malos. Ellos han permitido que sus departamentos de educación cierren escuelas, cambien a los directivos o entreguen la administración a otras instituciones. Otros estados han adoptado prácticas intermedias antes de remover a los directivos. Es el caso del departamento de educación de Carolina del Norte, que somete a los colegios con resultados muy deficientes a una asesoría obligatoria⁶. Por el lado de los estímulos, una medida muy difundida es entregar

⁶ El departamento interviene las escuelas con un equipo de tres a cinco expertos que se instalan en la escuela por un año. Entre ellos hay especialistas en currículo, en métodos didácticos y un administrador que ayuda a relocalizar recursos. Las 15 escuelas intervenidas

premios financieros por desempeño, como el Premio de Excelencia Académica otorgado en Chile.

La literatura sobre los sistemas de rendición de cuentas, evaluación e incentivos es vasta y compleja. Cada uno de los factores antes enumerados ha sido investigado en forma independiente y extensa, además los países los aplican en combinaciones diversas. Recopilar cuáles son los mecanismos que utilizan los distintos países para exigir resultados a sus educadores y alumnos, es un paso necesario para definir un marco de referencia que permita analizar la situación imperante en Chile y proponer un modelo que se adecue a nuestra realidad.

En este artículo sólo se abordará el tema de la rendición de cuentas de los alumnos. Se escoge este punto porque se lo ha estudiado menos y porque la evidencia apunta a que esta instancia puede tener efectos en el rendimiento y, a la vez, gatillar la exigencia de cuentas en otros niveles. Cuando se demanda alcanzar un estándar a los alumnos, los padres empiezan a exigir cuentas a los responsables de la gestión educativa de los establecimientos. Esto induce a los profesores y administradores a refocalizar sus energías hacia el logro de aprendizajes. La rendición de cuentas de los alumnos genera un círculo virtuoso de mejoramiento de la calidad.

II. LOS SISTEMAS DE RENDICIÓN DE CUENTAS DE LOS ALUMNOS

El alumno rinde cuentas cuando tiene que demostrar si ha aprendido lo que se espera de él según estándares prefijados y cuando debe atenerse a las consecuencias asociadas a los distintos niveles de logro. En la práctica, los sistemas de rendición de cuentas por parte de los alumnos consisten en exámenes externos al establecimiento, con consecuencias individualizadas que condicionan la promoción o certificación de un tramo educacional específico, o que determinan la asignación de los alumnos a distintos niveles y modalidades de estudio según su rendimiento, como también el ingreso a instituciones prestigiadas y el acceso a becas. Además, se considera como consecuencia el solo hecho de entregar a los padres un informe con los resultados de sus hijos en los exámenes externos.

lograron cambiar su calificación deficiente en un año. Esta experiencia difiere de las asesorías entregadas en el marco del programa chileno llamado "P-900". En el P-900 se entrega asesoría más esporádica y la escuela no está sometida a la presión de mejorar para evitar una sanción posterior. Los asesores tampoco se miden con los avances obtenidos, como es el caso en Carolina del Norte. Para una descripción de la experiencia de Carolina del Norte, véase el estudio de caso N° 3 aparecido en R. Phelps, *Why Testing Experts Hate Testing* (2002).

Las recomendaciones y notas se consideran parte de un sistema de rendición de cuentas cuando éstas responden a estándares externos. Por ejemplo, cuando son validadas en alguna instancia posterior que compromete el prestigio de aquellos profesores o colegios que recurren a la inflación de notas o a las recomendaciones. En California, los establecimientos de enseñanza básica tienen que recomendar, sobre la base de las notas, a los alumnos que deberían ser aceptados en el curso de álgebra en las escuelas de educación media. Al final del curso, los alumnos rinden un examen elaborado por la Universidad de Standford y se hace pública la relación entre el número de alumnos que cada escuela de enseñanza básica recomienda y el número de alumnos que efectivamente logran aprobar el examen final de esa asignatura. De este modo las notas en matemática y las recomendaciones del colegio básico se ajustan a un estándar, porque a la escuela le interesa aparecer con un alto número de postulantes que respondan. Algo parecido sucede con las recomendaciones que entregan los profesores de educación media de elite en Inglaterra. Si los alumnos recomendados por un establecimiento no se comportan a la altura de la universidad, queda comprometido el prestigio de esa institución. Lo cual tiene como consecuencia que las nuevas recomendaciones pierden peso y que el valor del currículo de los profesores que allí trabajan disminuya.

En los sistemas en que los profesores y las escuelas tienen amplios grados de libertad para asignar las notas, es más difícil que opere un régimen de “rendición de cuentas”. En el aislamiento de la profesión docente hay una tendencia a perder perspectiva de lo que se considera el rendimiento adecuado para una edad determinada, por lo tanto los profesores tienden a adecuar las notas a lo que han logrado enseñar, a lo que los niños han podido aprender y a las presiones que ejercen los padres y alumnos. Todo lo cual se aleja de la idea central de la rendición de cuentas, que es responder a estándares definidos de antemano.

2.1. Razones para exigir cuentas a los alumnos

La evidencia empírica muestra que los alumnos responden a los incentivos vinculados a los sistemas de rendición de cuentas. La explicación de base es simple: “el aprendizaje no es un acto pasivo; requiere tiempo y el compromiso activo de los que aprenden. Los alumnos tienen muchas alternativas para el uso de su tiempo, por lo tanto la intensidad del esfuerzo que realicen para aprender dependerá de que ellos consideren que los beneficios asociados al aprendizaje (tanto las recompensas intrínsecas

como extrínsecas del aprendizaje) sean superiores a los costos. Cuando los beneficios de aprender superan los costos, el esfuerzo de los alumnos aumenta y con ello el aprendizaje⁷. Los sistemas educativos exitosos toman en cuenta esta realidad haciendo evidentes *de antemano* las recompensas de estudiar y los costos de holgazanear.

Se podría pensar que la vida es la que mejor exigirá cuentas a los alumnos. Quienes no se aplican, probablemente tendrán acceso a trabajos menos interesantes o a instituciones de educación superior de menor nivel. Sin embargo, los costos de descuidar el estudio se experimentan tarde, cuando ya es difícil remediarlo. La mayoría de los niños y adolescentes no tienen la madurez cognitiva y psicológica para anticipar consecuencias distantes⁸. Por lo tanto, éstas no actúan como un incentivo al estudio. Es necesario entonces aproximar las consecuencias con un sistema de rendición de cuentas que presente a los jóvenes recompensas y sanciones *tempranas*.

Se supone que es responsabilidad de los padres el anticipar consecuencias a sus hijos, ya que son ellos quienes están más interesados en su futuro. Sin embargo, hay asimetrías de información entre el sistema educacional y los apoderados. Si éstos no tienen información que ponga en perspectiva cuánto saben sus hijos, cuánto saben sus pares y qué se espera que dominen en cada edad, difícilmente podrán ejercer este rol con un criterio de realidad. No basta con entregar información acerca del desempeño académico de la escuela y las notas de los alumnos. Si el establecimiento infla las notas que entrega a los alumnos, porque hace pruebas muy fáciles que todos pueden contestar, los padres podrían concluir erróneamente que su hijo está aprendiendo bien aunque la escuela no sea buena. Un sistema externo de rendición de cuentas que entrega datos de la escuela y de los alumnos individuales permite triangular dicha información, lo que fortalece el papel de los padres previsores que desean orientar bien a sus hijos.

Sin embargo, es necesario considerar que los adultos también tienen problemas con la anticipación de consecuencias futuras. Una muestra de ello es que los fondos de previsión obligan a cotizar en forma obligatoria. Esta actitud paternalista se sustenta en que es prácticamente imposible para la sociedad dejar que las personas experimenten directamente las conse-

⁷ John H. Bishop, "The Impact of Curriculum-Based External Examinations on School Priorities and Student Learning" (1996), p. 658. Para evidencia de los efectos de los exámenes externos, véase la sección sobre evidencia empírica de este artículo.

⁸ Para una explicación clínica de las dificultades de anticipación de consecuencias, véase la descripción del desarrollo de la función cognitiva de interiorización y de planificación de la conducta en R. Feuerstein, *Instrumental Enrichment: An Intervention Program for Cognitive Modifiability* (1980), pp. 97-99.

cuencias de su irresponsabilidad. Humanamente no se les podría dejar sin sustento si no han sido previsores. Los estudios psicológicos sobre el “optimismo injustificado” confirman la necesidad de apoyar a las personas para que sean previsoras. Éstos constatan que la mayoría de las personas piensan que se verán menos afectadas que el promedio por acontecimientos negativos en el futuro y que serán más propensas que el promedio a ser beneficiadas con hechos positivos⁹. La consecuencia directa de esta actitud es no invertir esfuerzo para prevenir posibles eventos negativos. En esta línea de argumentación, el adelantar las consecuencias de la falta de estudio facilitaría la corrección del optimismo injustificado, ya que los individuos aprenderían de la experiencia¹⁰.

Por otra parte, cuando el estado financia casi la totalidad de la educación de los niños y jóvenes es lícito exigirles cuentas. Al estado no le es indiferente si el alumno aprovecha o no las oportunidades que se le ofrecen ya que, por una parte, se estarían desperdiciando recursos y, por otra, los costos que genera una persona con educación deficiente los tiene que asumir luego la sociedad.

La rendición de cuentas por alumno contribuye a mejorar el rendimiento al cambiar la estructura de incentivos que opera al interior de los cursos. Las escuelas que tienen sistemas de notas relativas y las que permiten que los profesores las fijen generan una dinámica al interior de los cursos que incentiva el que todos ganen si nadie se esfuerza. Perder clases e interrumpirlas con indisciplina es un beneficio para los alumnos porque disminuye la materia que se acumula para las pruebas. Al mismo tiempo, cuando todos se ponen de acuerdo tácitamente para estudiar poco o no presentar trabajos ni tareas, fuerzan a los profesores a poner buenas notas por desempeños pobres¹¹. Los profesores no pueden reprobar a todos los

⁹ Véase N. D. Weinstein, “Unrealistic Optimism About Future Life Events” (1980), pp. 806-820.

¹⁰ La mayoría de las personas aprenden cuando tienen que enfrentar consecuencias negativas. Sin embargo, hay un número de optimistas injustificados que pueden tener dificultades para aprender de las experiencias de fracaso y mantener su actitud infundada frente a los nuevos desafíos que se les presenten. Pueden creer que aunque les fue mal una vez, después les puede ir mejor si tienen suerte, o por último que las consecuencias de no terminar los estudios o de no obtener un diploma o de no ingresar a buenas universidades no serán graves. Sin embargo, vale la pena establecer medidas que pueden ayudar a todos aquellos previsores que se benefician con información y a todos los optimistas injustificados que aprenden de la experiencia.

¹¹ De hecho en Chile hay evidencia de que los profesores han subido las notas sin un aumento de rendimiento concomitante. En los últimos 30 años, las notas han sufrido un proceso de inflación de un punto, pero esta alza no fue acompañada de una alza en el dominio de conocimientos en las pruebas de ingreso a la universidad. Véase G. Donoso, M. A. Bocchieri, E. Ávila, N. Núñez, et al., *El Sistema de Admisión: Orígenes y Evolución* (2000).

alumnos, en parte porque serían evaluados como malos pedagogos¹². Por lo tanto, se genera una cultura escolar en que los alumnos se “persuaden” unos a otros para imponer la solución cooperativa de no estudiar. Los alumnos ofrecen su amistad y respeto hacia los que tienen actitudes contrarias al estudio y denigran a los alumnos que demuestran interés. Este tipo de dinámicas ha sido ampliamente estudiado¹³. Chile no parece ajeno a esta realidad. Una profesora jefe de un liceo santiaguino manifiesta en una entrevista: “Cuando alguien entrega un trabajo lo pifian y se burlan de él. Algunos se avergüenzan y no vuelven a cumplir con las tareas para que no les hagan el vacío. A los que tienen más personalidad no les importa, menos mal”¹⁴. el autor (2000) constata que la denigración de los pares tiene un efecto negativo en el rendimiento. Utilizando la base de datos del NAEP¹⁵ de 1998, Johnson estudió la relación entre el rendimiento en lectura y la percepción de los alumnos respecto a la valoración del estudio entre sus compañeros. Encontró que en cuarto básico hay un fuerte efecto de la presión grupal cuando se controla género, raza, educación de los padres, nivel socioeconómico y recursos educativos en el hogar. Un alumno típico de cuarto básico disminuiría en un 8,5% su puntaje si percibe que “sus amigos se ríen de las personas que tratan de hacerlo bien en la escuela”. Las investigaciones demuestran que estas dinámicas se diluyen cuando los estándares se fijan externamente, ya que los alumnos constatan que no se obtiene nada a cambio de coludirse y que el costo de aceptar las normas del grupo tiene que sopesarse con las consecuencias que enfrentará al no cumplir con los estándares¹⁶.

¹² Obviamente, hay profesores y culturas escolares que logran entusiasmar a los alumnos e imponer un estándar por sí mismos.

¹³ Un ejemplo ilustrativo, en entrevistas realizadas a alumnos de los últimos años de enseñanza básica en Ithaca, Nueva York, los niños demostraron que habían internalizado ciertas normas para que no los consideraran “mateos”; entre ellas figuraban: evitar el contacto visual con el profesor, no entregar tareas anticipadamente para obtener créditos extras, no levantar frecuentemente la mano en clase, conversar y pasarse papelititos durante las lecciones, señalando así que prefieren tener buena reputación entre sus compañeros que entre sus profesores (J. H. Bishop, F. Mane y otros, “The Role of End of Course Exams and Minimum Competency Exams in Standards-Based Reforms”, 2001). Otro estudio también grafica el punto. Se trata de una investigación en nueve establecimientos de enseñanza media en California y Wisconsin, en la que se concluyó que menos del 5% de los alumnos se clasificaban a sí mismos en el grupo de los estudiosos. Estos alumnos eran los que estaban menos contentos con su grupo de pertenencia y cerca de la mitad deseaba pertenecer a otro grupo (L. Steinberg, B. Brown y S. M. Dornbusch, *Beyond the Classroom*, 1996, citado por Bishop, 1998).

¹⁴ M. Aguilar, “Tiempos Violentos: Se Cierra un Agitado Año Escolar”, *El Mercurio*, 14 de diciembre 2001.

¹⁵ El NAEP es el sistema nacional de evaluación del progreso educacional de Estados Unidos.

¹⁶ Véase J. Bishop, “The Impact of Curriculum-Based External Examinations on School Priorities and Student Learning” (1996).

Por otra parte, los sistemas que exigen cuentas a los alumnos permiten certificar de manera más objetiva las destrezas y conocimientos alcanzados por ellos, lo que a su vez facilita que el sistema laboral recompense tempranamente a aquellos alumnos que se han esforzado más en la escuela. Esta recompensa se constituye en una poderosa señal que induce al estudio a los que están en el sistema escolar. El sistema actual en Chile no permite diferenciar en forma fácil y confiable entre los alumnos que han obtenido su certificado de enseñanza media cumpliendo con éxito los programas y aquellos que lo han logrado por el sólo hecho de sentarse en clases. De hecho, los sueldos ni siquiera reconocen grandes diferencias entre los individuos que han cursado unos pocos cursos de enseñanza básica y los que casi han completado educación media, probablemente porque la diferencia en años de estudio no tiene una relación directa con las destrezas adquiridas por los alumnos (Beyer, 2000). En Estados Unidos, donde la aplicación de pruebas de competencias básicas de lectura, redacción y matemáticas fue prohibida por la Comisión de Igualdad de Oportunidades Laborales en los procesos de selección, la diferenciación de salarios por habilidades entre los alumnos recién egresados de la enseñanza media también es mínima (Bishop, Mane y otros, 2001). Sin embargo, en el transcurso de los años, los empleadores empiezan a distinguir en el trabajo quiénes saben y se esfuerzan más. A éstos los ascienden, les dan mayores posibilidades de capacitación y mejores sueldos. El resultado es que las empresas terminan recompensando tardíamente a quienes se habían esforzado más en la escuela. De hecho, las investigaciones demuestran que las correlaciones entre destrezas básicas evaluadas en enseñanza media y los sueldos de los individuos a la edad de treinta años son más altas que cuando se hace la misma correlación a los diecinueve años, manteniendo constante el número de años estudiados¹⁷. Con este sistema tardío de reconocimiento del esfuerzo escolar se pierde la señal que podría entregar el mercado laboral a los colegios. La certificación de destrezas implícita en todo sistema de rendición de cuentas de alumnos permitiría identificar y recompensar más tempranamente a los que estudian más, recuperándose así esta poderosa señal.

La teoría del principal-agente¹⁸ integra los puntos anteriores en un marco conceptual que permite explicar bien el aporte que hacen los exáme-

¹⁷ Véanse J. H. Bishop, F. Mane y otros, "The Role of End of Course Exams and Minimum Competency Exams in Standards-Based Reforms" (2001); John H. Bishop, "Impact of Academic Competency on Wages, Unemployment, and Job Performance" (1992); N. Grubb, "The Varied Economic Returns to Postsecondary Education", pp. 365-382, y H. Farber y R. Gibbons, "Learning and Wage Dynamics" (1996).

¹⁸ Por principal se entiende quien encarga un servicio y por agente quien lo realiza en nombre del principal.

nes externos al mejoramiento de la calidad de la educación. Bishop y Wößmann (2001, p. 5), lo elaboran en su artículo “Institutional Effects in a Simple Model of Educational Production”, en el que plantean que los exámenes externos con consecuencias individuales para los alumnos ayudan a romper el problema principal-agente presente en educación. La teoría de principal-agente se hace cargo de la idea de que las instituciones tienden a organizarse de un modo que no aseguran su eficiencia. Por el contrario, están llenas de conflictos de intereses, contratos incompletos y de incentivos que apuntan en dirección opuesta. En las instituciones, el principal tiene un contrato implícito o explícito para que un agente actúe de parte de él. El agente tiene intereses propios y al hacerse cargo de la ejecución de la tarea tiene ventajas de información que el principal no goza. El principal, al contar con menos conocimientos, experimenta dificultades para monitorear las acciones del agente, esto deja margen para que el agente actúe en forma oportunista, en el sentido de buscar su propio beneficio en desmedro de la tarea que le han encomendado. En educación no hay un solo principal y un solo agente sino una cadena de relaciones de esta naturaleza. Simplificando tenemos que los padres son el principal último que le encarga al Estado o a la escuela la educación de sus hijos. El Estado se la encarga a los sostenedores de municipalidades y privados, quienes a su vez se la delegan a los directores y éstos a los profesores. Cada uno de estos “contratos” están afectos a problemas de monitoreo¹⁹ y cada uno de estos agentes puede usar los recursos de los cuales dispone para promover sus intereses o para maximizar el aprendizaje.

A su vez, señalan Bishop y Wößmann (2001), los diferentes grupos de agentes educacionales persiguen intereses múltiples. Los profesores por un lado se sienten gratificados cuando sus alumnos progresan y por otro tienen un interés genuino en aumentar sus ingresos dada una carga de trabajo o en disminuir la carga de trabajo dado un ingreso fijo. Por ejemplo, pueden utilizar parte de la clase en actividades accesorias que les resultan más entretenidas, evitando tareas necesarias que les producen más estrés. De esta manera los profesores enfrentarán intereses que se contraponen. Los directores y administradores también experimentan conflictos, por ejemplo, entre enfrentar situaciones que producen tensiones o dejarlas pasar.

¹⁹ Se puede pensar que el director no tiene problemas de monitoreo con sus profesores. En estricto rigor no debiera tenerlos porque podría observar sus clases y revisar su trabajo; sin embargo, en la práctica esta tarea se facilita con las evaluaciones externas de rendimiento. Los profesores pueden fluctuar en su desempeño y normalmente un director tiene que supervisar un gran número de maestros; por lo tanto, se hace difícil el monitoreo sin ayuda de indicadores de desempeño objetivos.

La consecución de los intereses personales —según Bishop y Wößman (2001)— será más fácil o más difícil, dependiendo de la institucionalidad en que se encuentran. Si en el arreglo institucional el desempeño de los alumnos se puede constatar y observar, el mejoramiento de su rendimiento reportará mayores beneficios para los agentes. Si por el contrario ellos tienen amplios márgenes de libertad para disminuir su carga de trabajo, sin que se note, esto traerá consecuencias negativas para el aprendizaje. Los exámenes externos de rendimiento de los alumnos alteran de manera fundamental los incentivos que experimentan los agentes del sistema educacional, porque objetivan y hacen transparente el nivel de logro alcanzado, facilitan la comparación entre alumnos y escuelas y por lo mismo permiten distinguir entre responsabilidades individuales de los niños, de las escuelas y de los distritos. Cuando existen estos exámenes, los márgenes de libertad, de los profesores y de los establecimientos, para actuar de manera oportunista se reducen y los incentivos para emplear el tiempo y los recursos de manera más eficiente aumentan. A través de la difusión de los resultados de los exámenes, los padres y cada uno de los principales dejan de estar en una abierta desventaja de información respecto de los agentes a los cuales se les ha encargado la educación. De esta manera, los principales pueden pedir cuentas a los agentes, exigiendo el cumplimiento de sus contratos. Los padres pueden evaluar el rendimiento de sus hijos, profesores y escuela; el director puede evaluar el rendimiento de sus profesores, los sostenedores a sus directores y el gobierno a los sostenedores. De esta manera se soluciona uno de los principales problemas de la relación principal-agente, lo cual explica, en parte, el efecto positivo de los exámenes externos con costos para los alumnos.

Antes de exponer la evidencia empírica al respecto, analizaremos otra serie de razones que fundamentan la instauración de un sistema de exámenes externos con costos para los alumnos. Éstas las hemos agrupado separadamente, ya que tienen un carácter más pedagógico.

2.2. Razones pedagógicas para instaurar exámenes de evaluación externos con consecuencias individuales

Los exámenes externos individualizados obligan a los establecimientos a concebir sus responsabilidades en términos menos abstractos que los porcentajes generales de logro de contenidos. Cuando se trata de niños específicos que no son promovidos o que no reciben su certificación, se personalizan las deficiencias. Acomodarse al hecho de que se ha perjudica-

do a individuos es difícil, por lo tanto, aumentan los incentivos psicológicos para que las escuelas y los profesores busquen mejores métodos para que más niños aprendan. Las investigaciones sobre los efectos de la implementación del sistema de exámenes externos en el estado de Texas afirman que éstas han generado una cultura de altas expectativas, sentido del desafío y entusiasmo para alcanzar las metas. La cual se ha acompañado de inmediatos esfuerzos para ayudar a los alumnos que tienen un desempeño pobre, ya que se ha instituido un sistema que permite identificar los problemas en forma temprana, evitando que se acumulen y se hagan difíciles de abordar²⁰.

Por otra parte, los estándares claros que proveen estas pruebas benefician directamente a los alumnos y a sus padres, ya que ellos saben qué se espera que aprendan y hacia qué tienen que dedicar sus esfuerzos. Entre otros, los sistemas de examinación de California, Florida, Alberta (Canadá) y los de Cambridge (Inglaterra) proveen señales claras sobre lo que los alumnos deben dominar. Del mismo modo, los exámenes externos obligan a los colegios a focalizar las energías en torno a lo que ellos evalúan. Las escuelas tienden a asumir múltiples tareas, por lo que los exámenes externos ayudan a centrar las prioridades. Mientras no se logre lo mínimo exigido, la escuela no podrá desviar energías a otros objetivos. En este sentido ayuda a aunar criterios, a aclarar los objetivos y a definir el estándar de la escuela. Puntos que son requisito para una buena gestión escolar²¹. Las pruebas también son una buena herramienta para elevar los estándares cuando las metas implícitas que imperan en la cultura escolar son bajas. En el caso de Chile, en que los resultados en las pruebas internacionales nos han mostrado repetidamente que el nivel educacional alcanzado es deficiente, nos beneficiaríamos con la posibilidad de elevar el estándar al cual aspiramos, ya que, en general, los profesores tienen una percepción que sobrevalora nuestra educación. Sólo el 5,7% de los directores y el 13,2% de los profesores calificó de mala la calidad de la educación del país²². Modificar las metas a las cuales aspiran los profesores es importante porque los estándares altos contribuyen a mejorar el rendimiento de los alumnos²³.

²⁰ Véase R. P. Phelps, "Test Bashing", Part 3: The Education Press's Cop-Out on Student Testing (2002).

²¹ Véase M. S. Tucker y J. B. Coddling, *Standards for Our Schools. How to Set Them, Measure Them, and Reach Them* (1998).

²² CIDE, "II Encuesta Nacional a los Actores del Sistema Educativo 2000" (2001).

²³ Véase N. D. Figlio, "Do High Grading Standards Affect Student Performance?" (2000). El estudio encuentra que los profesores que tienen estándares altos al evaluar tienden a tener un impacto positivo en el rendimiento en matemática y lenguaje. Comprueba que los niños de rendimiento alto se benefician más de los estándares altos cuando se encuentran en

Someter a los alumnos a un sistema de estándares fijos, lo que está implícito en un sistema de evaluación externa con repercusiones para los alumnos, quizás aparezca como una imposición rígida que puede acarrear costos emocionales a los niños. Sin embargo, exigencias proporcionales y similares a las que enfrentarán de adultos son válidas porque es importante que los alumnos generen la capacidad de lidiar con ellas. Por una parte, la mayoría de los niños poseen la capacidad para cumplir con los requerimientos escolares si se aplican a ello y si los profesores les dan la oportunidad de aprenderlos. Un niño de inteligencia normal (teóricamente cerca del 70% de la población) puede enfrentar los requerimientos de la enseñanza básica sin problemas. El 20% de los restantes podrá hacerlo con mayor dificultad, probablemente con reforzamiento pedagógico o tomándose un período más largo. Teóricamente el 10% tendrá necesidad de educación especial ya sea integrada o en escuelas diferenciadas²⁴. En educación media, probablemente estas cifras difieren según el tipo de educación escogida. Las vertientes académicas exigentes requieren mayores niveles de abstracción, lo cual implica que los porcentajes de la cohorte de edad que pueden pasar los exámenes académicos avanzados finales de sus respectivos países son menores. En Francia, por ejemplo, lo logra el 32%, en Alemania el 36%, en Israel el 40% y en Japón el 33%. Sin embargo, un mayor porcentaje puede alcanzar un nivel aceptable de conocimientos generales, equivalente al plan común de nuestro currículo de enseñanza me-

cursos de bajo rendimiento, y los niños con bajo rendimiento se benefician más con los estándares altos cuando están en cursos con alto rendimiento. Además se observa que los altos estándares reducen los problemas disciplinarios en la escuela. Kulik (1984), Monk (1994) y Bishop (1996) muestran que se aumenta el nivel de aprendizaje cuando los alumnos toman cursos más demandantes, aun cuando se controla por la calificación del profesor, por la habilidad y por el status socioeconómico del estudiante.

²⁴ Este 10% de niños con necesidades especiales es una cifra que se acepta internacionalmente. El 70% considerado en el rango normal incluye a los alumnos talentosos hasta los individuos con CI de 95 que se consideran de inteligencia normal. En el 20% que necesita ayuda pedagógica se incluyen los alumnos con inteligencia normal lenta y aquellos que tienen problemas específicos de aprendizaje como disfasias, dislexias, etc. Véase H. J. Eysenk, *A New Look to Intelligence* (1998). En Chile, en base al dato del 10% internacionalmente aceptado, se estima que el número total de niños en edad escolar que tienen necesidades especiales son cerca de 330.000. De ellos sólo 50.000 son atendidos por el sistema escolar, 46.049 en escuelas especiales y 5.500 integrados en establecimientos regulares. El resto no asisten a la escuela o bien asisten sin estar identificados como niños con necesidades especiales. De esto se desprende que un alto porcentaje de los alumnos en el sistema regular de enseñanza básica pueden, teóricamente, cumplir con los mínimos requeridos. Sin embargo, un sistema de exámenes de competencias mínimas debe permitir eximir al menos al 5% de la población matriculada para que pueda realizarse la integración de los menores con necesidades especiales, ya que se sabe que esto trae beneficios para su desarrollo intelectual y emocional. Véanse datos y una discusión del tema en Unicef, "Inclusión de Niños con Discapacidad en la Escuela Regular" (2001).

dia. Los países que han instaurado exámenes de competencias mínimas y pruebas de excelencia toman en consideración esta realidad, para lo cual diseñan evaluaciones de buen nivel para planes diferenciados de estudio. Así, en Francia encontramos las pruebas de bachillerato académicas y vocacionales. En el bachillerato académico los alumnos pueden escoger entre tres líneas: la literaria, la científica y la económica-ciencias sociales. Aunque todos tienen que demostrar dominio en la lengua materna, la elección de líneas deja espacio para individuos de diversos talentos. En el área vocacional hay 30 líneas²⁵.

Detrás de la aprensión de exigir a los alumnos el logro de determinados mínimos está la teoría de que al momento de rendir pesa más la habilidad que el esfuerzo. Ésta no es la idea que predomina en las culturas asiáticas que alcanzan altos niveles de logro ni en las experiencias que han tenido éxito con niños de minorías en Estados Unidos. Por el contrario, en ellas prevalece la creencia de que el esfuerzo y el trabajo duro son centrales para conseguir el éxito²⁶.

Los exámenes externos obligan a los alumnos a esforzarse y lograr determinados mínimos en cada ciclo. Esto trae importantes beneficios pedagógicos para los alumnos y profesores. Enfrentar las exigencias de un curso teniendo los prerrequisitos logrados facilita la tarea del niño y contribuye a sus posibilidades de éxito, lo cual aumenta su autoestima. También le allana el camino al profesor, ya que puede avanzar más rápido porque no tiene que volver hacia lo más elemental y puede trabajar con cursos que tienen variaciones individuales manejables. En Chile hay evidencia de que alrededor de 40% de los niños en pobreza no logra los conocimientos mínimos en lenguaje y matemática al terminar segundo básico²⁷. Seguramente ellos tendrán dificultades para abordar las materias de tercero, pues no dominan la lectura y las nociones básicas de número requeridas para avanzar en sus estudios. Los textos están pensados para niños que ya leen, por lo tanto quedarán desfasados para ellos y, así, se irán sumando las dificultades.

Otro importante motivo para defender los exámenes externos individuales es que éstos mejoran las relaciones entre los profesores y los alumnos, ya que los maestros quedan liberados de fijar los estándares para su curso. El profesor es visto como un entrenador que ayuda a superar la valla

²⁵ Véase E. D. Britton y S. A. Raizen, *Examining the Examinations: An International Comparison of Science and Mathematics Examinations for College Bound Students* (1996).

²⁶ Véase H. Stevenson y J. Stigler, *The Learning Gap: Why Our Schools Are Failing and What We Can Learn from Japanese and Chinese Education* (1992). Para una experiencia educativa exitosa en sectores desventajados socialmente, basada en el esfuerzo de los niños, estúdiese la política del programa educativo "Knowledge is Power Program (KIPP)" diseñado por Michael Feinberg.

²⁷ Véase B. Eyzaguirre y C. Le Foulon, "La Calidad de la Educación en Cifras" (2001).

más que como el que la impone. En los sistemas en que los alumnos y los profesores enfrentan consecuencias asociadas al rendimiento de los alumnos surge una comunidad de intereses y se evita la tensión entre dar apoyo al alumno y evaluarlo. En Irlanda se propuso cambiar los exámenes externos por pruebas administradas por los mismos profesores. El sindicato de los profesores de enseñanza media se opuso terminantemente a esta medida argumentando que suprimir los exámenes externos “produciría un distanciamiento automático entre los profesores, los pupilos y los padres. También abriría la puertas a posibles distorsiones provocadas por la presión de los padres...” Sostuvieron además que una de las mayores fortalezas del sistema irlandés ha sido la percepción de los estudiantes de que sus profesores son “abogados en relación con los exámenes de certificación nacionales en vez de sus jueces”²⁸. Algo así sucedía en Chile con los exámenes válidos. Los docentes de colegios particulares preparaban a sus estudiantes para el examen y aquéllos se sentían tan evaluados como éstos cuando lo rendían. El profesor presentaba al joven con una nota y no era bien visto que el rendimiento en la prueba difiriera significativamente de esta calificación. Había una simetría de intereses, el alumno necesitaba aprobar y el profesor deseaba demostrar que había enseñado bien. Una identificación entre los dos bandos redundaba en más tiempo dedicado al estudio. Además, esta cooperación entre profesores y alumnos puede ayudar a aumentar la motivación de los profesores. Preparar una clase que será recibida con una buena disposición hacia el aprendizaje es altamente reforzador, en cambio la apatía reiterada de los alumnos puede resultar muy desmotivador.

III. IMPACTO DE LOS SISTEMAS DE EXÁMENES EXTERNOS: EXÁMENES DE COMPETENCIAS MÍNIMAS Y EXÁMENES FINALES EXTERNOS BASADOS EN EL CURRÍCULO

Distinguir entre los exámenes de competencias mínimas (ECM) y los exámenes finales externos basados en el currículo (EEBC) es relevante porque cada uno de estos sistemas tiene un impacto diferente. Si bien un gran número de países y regiones los utilizan en forma combinada, un análisis preliminar independiente permite comprender mejor la evidencia empírica que existe al respecto.

²⁸ Citado en J. H. Bishop y F. Mane, “The Impact of Minimum Competency Exam Graduation Requirements on College Attendance and Early Labor Market Success of Minority Students” (1998).

3.1. Los exámenes de competencias mínimas (ECM)

Los exámenes externos de competencias mínimas definen un estándar de aprendizaje, el cual debe ser logrado por los alumnos. Como referencia, en Estados Unidos todos los estados tendrán funcionando cabalmente, cerca del año 2005, un sistema de certificación de la enseñanza media que requiere aprobar los ECM. Actualmente 18 estados retienen los diplomas de graduación si los alumnos no alcanzan los estándares esperados y 8 estados basan la promoción al curso siguiente principalmente en los resultados de los ECM (Carolina del Norte, Carolina del Sur, Delaware, Georgia, Louisiana, New Mexico, Ohio, Texas). Chicago exige asistir a cursos remediales en el verano a los alumnos que no aprueban los mínimos esperados²⁹.

En los estados con mayor proporción de población latina y negra los ECM son una práctica más difundida; el 79% de las minorías las tiene como requerimiento. Esta política es concordante con la investigación de Bárbara Lerner (1990), que sostiene que los ECM es una de las medidas que han tenido más impacto en el mejoramiento del rendimiento de los alumnos desventajados. Probablemente los exámenes de competencias mínimas ayudan a elevar el “piso”.

Si se revisa el conjunto de los estudios empíricos sobre los ECM se encuentran efectos positivos. Sin embargo, la magnitud del efecto resulta discutible. Cada una de las investigaciones incluye especificaciones diferentes en sus análisis estadísticos, por lo tanto, no es fácil establecer relaciones directas entre ellas. En esta exposición no se pretende determinar la dimensión del impacto, sólo se busca plantear el estado del debate.

Para constatar el efecto de los exámenes de competencias mínimas es necesario que las investigaciones tomen en cuenta las condiciones básicas que se deben cumplir para que éstos tengan efecto y eviten posibles errores de medición. Pasar por alto estos puntos puede llevar a una mala interpretación de los resultados, atribuyendo efectos inexistentes o negando una posible contribución. A continuación revisaremos primero los factores que teóricamente contribuyen a que los ECM tengan efecto y luego los posibles elementos que se deben considerar para no incurrir en errores de medición. Luego analizaremos las investigaciones a la luz de estos puntos.

- Se debe esperar un efecto mayor de los ECM cuando se asocian a “costos” altos. Por lo tanto, *las investigaciones debieran diferenciar entre aplicaciones con altos o bajos costos (high stakes vs. low stake)*. Ejemplo de costo alto es retener la certificación de enseñanza media si no se pasa el

²⁹ NGA Center for Best Practices, “Graduation Exit Exams” (datos al 2001).

examen, promediar con ponderación alta la nota del examen con el resto de las calificaciones o exigir asistencia a un curso de verano. Una consecuencia de bajo costo puede ser el entregar los resultados al establecimiento.

- En muchos estados, los sistemas de examinación tienen un período de gracia antes de aplicar las sanciones y premios. Éstos se instituyen para dar un tiempo de preparación, de modo que las escuelas y los alumnos puedan ajustar sus esfuerzos a las nuevas metas. *Las investigaciones no debieran darles el mismo status a los casos en que se están aplicando las consecuencias que a aquellos en que son sólo una amenaza potencial.* Por ejemplo, la política de exigir un examen al final del décimo año en Massachusetts se anunció el año 1993, se aplicó por primera vez en 1998 y sólo entrarán en vigencia las consecuencias asociadas a ella el año 2003.

- La evidencia señala que *no se deben comparar en igualdad de condiciones a los alumnos que están bajo regímenes de ECM exigidos por escuelas, por el distrito o por una autoridad central.* Si los alumnos pueden evadir los requerimientos de un distrito o escuela cambiándose de establecimiento, tendremos que los resultados observados en los distritos que tienen exámenes responderán a un sesgo de autoselección de los alumnos más que a un aumento del estudio. Por otra parte, cuando los exámenes son descentralizados y demasiado numerosos, puede surgir una confusión con respecto al valor de cada uno de ellos. En consecuencia, la información se hace más difícil de interpretar por el mundo laboral. Al no poder distinguir, los empleadores tenderán a no diferenciar salarios, disminuyéndose los incentivos asociados a la aprobación de los exámenes. Se concluye entonces que no se debiera comparar en igualdad de condiciones los sistemas que aplican pruebas muy localizadas con los que aplican pruebas más centralizadas.

- No se debe esperar el mismo efecto de aquellos sistemas que imponen puntajes de corte bajos para aprobar sus exámenes y de aquellos que ponen la vara más alta. Efectivamente, al menos en Estados Unidos hay estados que aplican criterios extremadamente fáciles de superar y otros que son más exigentes. Los índices de reprobación de la primera vuelta del examen de egreso de educación media varían de 7% en Mississippi, 27% en New Jersey, 30% en Tennessee, 34% en Virginia, a 46% en Texas³⁰. Cuando la meta por alcanzar es muy baja, ésta no actúa como incentivo al

³⁰ Se puede pensar que estas variaciones no se deben a diferencias en el nivel de dificultad donde se fija el punto de corte, sino que se refieren a diferencias en el nivel de dominio de los contenidos examinados. Sin embargo, el orden de magnitud sugiere que están presentes ambos factores.

esfuerzo, por lo tanto el aumento del rendimiento debiera ser casi inexistente. Los países y estados también difieren entre sí en el número de veces que permiten repetir el examen; si éste se puede repetir muchas veces el incentivo al estudio será menor para la primera vuelta. En las localidades en que quienes fijan el estándar son igualitaristas, la tendencia es imponer mínimos extremadamente bajos para que nadie repruebe los exámenes. En estos casos, los ECM podrían bajar los estándares impuestos por los propios profesores en sus clases, produciendo un efecto nocivo en el aprendizaje. Lo mismo sucede cuando quienes fijan el estándar valoran altamente el tiempo libre para los alumnos³¹.

- Difusión de la existencia de los exámenes de competencias mínimas: es obvio que si los alumnos desconocen que serán evaluados no harán esfuerzos extraordinarios para salir bien en las pruebas. Caterell (1994) constató que en cuatro estados en que realizó un estudio sobre los ECM, la mitad de los alumnos de enseñanza media no sabían de la existencia del requisito de aprobar un examen para graduarse. Cerca del 45% sabía de este requisito en 9° año y este porcentaje subía levemente a 58% en el 11° año. Más grave resulta este dato si se considera que las pruebas se podían tomar desde 9° y 10° año y que ya las habían rendido más de la mitad de los alumnos. Por lo tanto, *debe tomarse en cuenta el grado de conocimiento de la existencia del examen al juzgar la efectividad de los ECM*³².

- La evidencia indica que se potencian los efectos de los ECM cuando los alumnos cuentan con información detallada acerca de lo que se espera de ellos en las pruebas y tienen buenos materiales didácticos para el estudio independiente. En Inglaterra, el *syllabus* de los exámenes es preciso y los estudiantes pueden encontrar una diversidad de libros que los ayudan a estudiar para las pruebas de certificación de cada ciclo. *Para ponderar los efectos de los ECM se debiera tomar en cuenta si el sistema está diseñado de manera tal que los alumnos pueden encontrar ayuda para hacer fructificar sus esfuerzos.*

- Interacción entre el rendimiento de los alumnos y los exámenes de competencias mínimas. Teóricamente, *los efectos de los ECM son distintos*

³¹ Véase R. M. Costrell, *A Simple Model of Educational Standards* (1994). Para solucionar el problema de los gobiernos igualitaristas que fijan estándares mínimos muy bajos propone evitar los criterios dicotómicos de aprobar/reprobar. También sugiere adoptar categorías múltiples para que los alumnos perciban que cada esfuerzo marginal que realicen se reconocerá en su certificado.

³² Es factible que las pruebas fuesen excesivamente fáciles para una mayoría de los alumnos y por eso no se percataron de su existencia. Otra posibilidad es que el certificado de graduación carecía de valor para ellos.

según el rendimiento previo de los alumnos: debieran ser escasos o nulos entre aquellos alumnos de alto rendimiento que se preparan para exámenes avanzados, ya que no tendrían que hacer ningún esfuerzo especial para pasar las exigencias de los exámenes de competencias mínimas. En cambio los alumnos de rendimiento bajo y promedio se verían más afectados, porque tendrían que hacer un esfuerzo extra para aprobarlos. Para los que están en el extremo inferior de la curva de rendimiento, el costo del esfuerzo puede superar los costos asociados a no aprobar, lo cual llevaría a que los alumnos abandonen todo intento de estudio.

- Efectos diferenciados según la edad. Al parecer los efectos de los ECM son mayores cuando los alumnos están empezando su escolaridad que cuando la están terminando. En los cursos de básica es más fácil remontar la brecha entre lo que el niño ha logrado y lo que se espera de él, porque los déficits son menores que en años posteriores, cuando éstos se han acumulado período tras período. A esa edad la meta se puede superar con un esfuerzo menor de parte de los profesores y alumnos. Por lo tanto, es más fácil que respondan a los incentivos provistos por los ECM. En cambio, en la enseñanza media la acumulación de déficits puede ser tal que los incentivos no alcancen a compensar el esfuerzo que hay que realizar.

- Repercusiones de la deserción y la repitencia. Las investigaciones debieran hacerse cargo de controlar si el aumento del rendimiento es real o si se debe a que en la población de los alumnos matriculados se ha producido un “descreme” porque han desertado los que tenían mal rendimiento y sólo se han quedado los mejores. También hay que tomar en cuenta si el aumento del rendimiento se produce sólo por un factor de sobreedad de los alumnos repitentes, en cuyo caso habría que ponderar el costo-efectividad de la medida.

- Inflación de los resultados por familiaridad con las pruebas. En el análisis de los efectos de los exámenes externos, interesa saber si los aumentos de puntaje son un reflejo real de mayores grados de aprendizaje o si sólo son resultado de un adiestramiento en la mecánica y estilo del examen. No es lo mismo analizar ganancias en los exámenes que normalmente se les aplican a los alumnos que en exámenes muestrales diseñados para investigar.

- Presencia de otras medidas que podrían estar afectando al rendimiento. Es frecuente que en los estados que han introducido ECM, también se apliquen otras medidas como la incorporación de nuevos recursos mone-

tarios y pedagógicos, igualación de presupuestos entre las escuelas, disminución de alumnos por curso, etc. No tomar estas variables en consideración puede llevar a la conclusión errada de que son los exámenes los que producen los efectos cuando en realidad son otro tipo de variables las que explican las ganancias. Sin embargo, hay otra serie de políticas que deberían acompañar a los ECM, porque su ausencia crearía una situación unilateral de exigencia hacia ellos que sería abusiva e injusta. Por ejemplo, la introducción de medidas de rendición de cuentas a las escuelas y profesores no son independientes del funcionamiento de los ECM. También es necesario distinguir los problemas de endogeneidad que se presentan. Es posible que los ECM tengan como consecuencia un aumento de los salarios de los profesores y un aumento en el gasto en educación. En Estados Unidos los padres estarán más dispuestos a aprobar presupuestos más altos para las escuelas de su distrito si perciben que es necesario mejorar la educación de sus hijos. Por lo tanto, sería errado controlar por salario del profesor y por otras variables que no son independientes de los ECM.

- Es necesario tomar en cuenta la posibilidad de que los resultados estén sesgados por la calidad de la educación previa a la instauración de los ECM. Puede darse el caso de que sólo aquellos estados que tienen una buena calidad de educación y que se preocupan de ella sean los que imponen exámenes de este tipo. En este caso, si se realiza un estudio de correlación simple entre presencia de las pruebas y resultados puede concluirse equivocadamente que tienen un efecto positivo. También está la posibilidad inversa, de que sólo impongan exámenes aquellos estados que tienen peores resultados previos, como una manera de mejorar rápido, con lo cual se sesgarían los resultados contra los ECM. Este punto es importante porque algunas de las críticas a los ECM en EE.UU. surgen de estudios de correlación simple que afirman que dichas pruebas no tienen efecto. Sin embargo, al analizar el contexto, se constata que los estados que las aplican son los que tienen una situación inicial más precaria.

Para hacerse cargo de este problema algunas investigaciones comparan resultados de un mismo estado pre y postimplementación de los ECM. Otras han propuesto controlar por la calidad de la educación previa de los alumnos. Calculan cuántos puntos de rendimiento avanzan los estudiantes desde el grado 8° hasta el 12° en los estados con ECM y lo comparan con el avance que realizan los alumnos que no están sometido a este tipo de exigencias. Suponen que los ECM influyen en el rendimiento de los alumnos con algunos años de anterioridad a la aplicación del examen, ya que profesores y alumnos empiezan a invertir mayores esfuerzos con anteriori-

dad. Esta preparación temprana sería un efecto directo de los ECM. Sin embargo, si se observan diferencias de rendimiento previas al período en que se estima que inciden los ECM, no se podrían atribuir las diferencias de rendimiento a los ECM sino que a la calidad de la educación general de la región. Sin embargo, es discutible la determinación del grado en que empiezan a hacer efecto los ECM.

- Presencia de otros factores que tienen incidencia en el rendimiento. Hay una serie de factores que la mayoría de las investigaciones controlan porque ya se sabe o se intuye que tienen incidencia en el rendimiento de los alumnos; entre otros, nos referimos al nivel socioeconómico y educativo de los padres, recursos educativos en el hogar, gasto por alumno, presencia de cultura asiática, etc.

A continuación se presentan las investigaciones empíricas que se han realizado en torno a los ECM. En general los estudios toman en cuenta algunos de estos factores y obvian otros; sin embargo, el conjunto de ellos da indicios de que los exámenes externos con consecuencias para los alumnos incentivan el esfuerzo y con ello el aprendizaje.

A. Estudios que comparan el rendimiento pre y postinstauración de ECM en distintos estados de Estados Unidos

Roderick³³ estudia las ganancias de aprendizaje entre el período anterior y posterior a la puesta en marcha del sistema de ECM en Chicago y observa que esta medida tuvo efectos positivos en el rendimiento de los alumnos. En este estado se exige aprobar un mínimo para ser promovido en 3º, 6º, 8º y 9º. Los alumnos que no aprueban los exámenes de matemática y lenguaje tienen que quedarse a un curso de verano, al finalizarlo tienen que repetir el examen y son promovidos al curso siguiente sólo si alcanzan el estándar exigido. En esta investigación el porcentaje de alumnos de peor rendimiento inicial que logra, en el primer intento, el puntaje de corte exigido aumenta 11 puntos porcentuales en el período que se impone el sistema. Entre los alumnos que tienen un retraso pedagógico inicial moderado también se observa un aumento en el número de los que logran superar el examen: éste es de 7 puntos porcentuales para el primer ensayo y alrededor de 30 puntos porcentuales en el segundo intento. Es interesante notar que los alumnos que asisten a las clases de verano logran mejorar su rendi-

³³ M. Roderick, *Ending Social Promotion: Results from the First Two Years* (1999), citado en J. R. Betts y R. M. Costrell, "Incentives and Equity under Standard Based Reform" (2001).

miento y que esta mejoría se mantiene a lo largo del año siguiente; sin embargo los niños vuelven a quedarse atrás frente a los nuevos aprendizajes. Todo indica que son niños que necesitan más reforzamiento y tiempo para aprender. Esta investigación es importante porque aporta evidencia que contradice la posición teórica de que los alumnos que están inicialmente muy lejos de la meta no responderán a los incentivos porque es demasiado el esfuerzo implicado. Como se observó, los alumnos que tenían peor rendimiento respondieron activamente a la imposición del estándar. En esta investigación y en las tres siguientes no se puede controlar si los efectos encontrados son atribuibles a los ECM o a una serie de otras medidas que se implementaron en esta ciudad o en los respectivos estados para mejorar la calidad de la educación.

Deere y Strayer (2001) investigan los efectos del sistema de rendición de cuentas en Texas, el cual evalúa a sus alumnos todos los años en 4°, 8° y 10°, entrega resultados individuales a los padres y elabora informes de las escuelas con el porcentaje que pasan o fallan las pruebas, calificando a los establecimientos en cuatro categorías: académicamente inaceptables, aceptables, reconocidas y ejemplares³⁴. Estos resultados tienen alta visibilidad pública y los superintendentes pueden decidir no renovar los contratos de los superintendentes de área, directores de escuelas y sus subdirectores académicos si no mejoran sus resultados. Los incentivos para los alumnos en estas pruebas se pueden considerar medianos, ya que todavía no entra en vigencia la retención de aquellos alumnos que no pasan las pruebas. Sólo se condiciona la certificación de los alumnos al final de la enseñanza media y se entregan los resultados a los padres. Además, los puntajes de grado a grado son comparables, de manera que si un alumno tiene un puntaje aceptable en los cursos iniciales, obtiene la señal de que si sigue por ese camino logrará aprobar el examen de graduación. Desde 1994 a 1999 la fracción de alumnos que pasaron la prueba en matemática aumentó de 58% a 85% y en lectura de 74% a 86%. Lo interesante de notar es que sólo se observó un aumento tan significativo en las pruebas que estaban incluidas en el sistema de calificación de las escuelas (matemática y lectura), aun cuando los resultados de las otras dos pruebas (las de ciencias) se publican cada vez. En 8° las diferencias fueron de 32 puntos porcentuales en matemática, 16 en lectura, 12 en ciencias y 6 en ciencias sociales. En esta investigación se comprueba también que los efectos positivos que tiene la medida de los

³⁴ Para establecer esta categoría también se consideran indicadores de deserción, asistencia. No se incluyó un indicador de nivel socioeconómico porque las pruebas no son relativas sino que referidas a criterio. Por lo tanto, se estima que todos los niños, dados los recursos que se entregan a las escuelas, pueden lograr los objetivos fijados.

ECM fueron más fuertes en los alumnos con habilidades iniciales que están bajo el nivel de dominio exigido. Esto se puede explicar con dos razones. La primera es que la vara impuesta no constituyó un desafío para los alumnos con buen rendimiento inicial, y la segunda es que los alumnos con problemas fueron los que recibieron más recursos y atención de parte de los profesores, ya que el estado permite decidir el uso de recursos en los establecimientos. De todo esto se concluye que, como es de esperar, las escuelas públicas de Texas y sus alumnos responden donde están los mayores incentivos. Esta investigación no controla por el efecto que tiene el “entrenamiento para la prueba”. Tampoco lo hace la investigación en Chicago ni la de Hess y Lockwood³⁵, quienes también encuentran evidencia de una influencia positiva en aplicaciones pre y post de los ECM en Alabama. Sin embargo, la investigación presentada a continuación controla esta variable y concluye que el logro no se debe al “entrenamiento para la prueba”.

Betts y Costrell (2001) examinan los resultados del sistema impuesto en Massachusetts y encuentran un aumento sobre 5 percentiles aun en la etapa inicial en que no han entrado en vigencia las consecuencias de reprobación los exámenes. El estudio compara el rendimiento de la generación de 4° básico anterior a la aplicación de las pruebas y de la posterior. Se controló por el rendimiento previo de los alumnos en cursos de 3° básico para asegurar que el efecto no estuviera sesgado por habilidades diferenciales de cada curso. Además se tomó una prueba de rendimiento anexa que medía conocimientos y destrezas para controlar el “entrenamiento para la prueba” y se encontró un aumento de 8 percentiles, lo que refleja que los aprendizajes logrados son sólidos y van más allá de un mero entrenamiento en los ítems de la prueba. Parte importante de la consistencia en el logro se puede atribuir a la alta calidad de los exámenes. Fueron elaboradas con estricto apego al currículo e incluyen un número importante de preguntas abiertas, otras de ensayo además de las tradicionales preguntas de alternativas.

Fredericksen (1994) realiza un estudio de comparación pre y post considerando una muestra más grande e intenta distinguir efectos diferenciados de los ECM rendidos en el 12° grado con consecuencias fuertes, moderadas y débiles. La variable dependiente es el rendimiento en el NAEP³⁶ de los alumnos de 9 años de la generación de 1978 sin ECM y de

³⁵ A. C. Hess y R. E. Lockwood, “The Relation of a Basic Competency Education Program to Overall Student Achievement: A State Perspective” (National Council on Measurement in Education Annual Meeting, San Francisco, 1986), citado en B. A. Jacob, *Getting Tough? The Impact of High School Graduation Exams* (2000).

³⁶ NAEP (National Assessment of Educational Progress) es una prueba nacional de rendimiento de Estados Unidos, a la que se someten voluntariamente la mayoría de los estados.

aquellos de la generación de 1986 con ECM en 27 estados. Para esta edad encuentra un aumento considerablemente mayor en los estados con ECM fuertes, que en los alumnos de similares características con ECM moderados o débiles. Repitió el diseño de la investigación con las cohortes de 13 y 17 años. Observó que el efecto se mantenía en las cohortes de 13 años y que en las de los alumnos de 17 años sigue encontrándose un efecto positivo pero menor que en los casos anteriores. Este fenómeno es curioso ya que a esta edad la consecuencia severa se refería a la certificación de enseñanza media. Fredericksen nota que en todos los segmentos de edad en que hay efecto el aumento de puntajes fue mayor en los ítemes que medían destrezas de nivel superior, comparado con los ítemes que implicaban respuestas simples y rutinarias. Esta investigación controla por el entrenamiento para la prueba porque el NAEP no es la prueba que aplican normalmente los estados.

B. Estudios de corte transversal con muestras nacionales en EE.UU.

Winfield (1990)³⁷ utilizó la muestra nacional del NAEP entre los años 1983-84 para comparar el rendimiento en lectura de los estados que aplican pruebas de ECM con los que no lo utilizan. Controló por una cantidad de variables individuales, de escuela y regionales. Encontró que los alumnos en sistemas con ECM tenían mejor rendimiento que sus pares que no rendían este tipo de pruebas, y encontró un efecto positivo aún mayor entre los alumnos hispánicos y afroamericanos.

Bishop, Mane y otros (2001) estudiaron el efecto de los ECM, tomando como variable dependiente el promedio de los estados en los resultados de las pruebas del NAEP de 1992, 1994, 1996 y 1998. En 4° y 8° básico estimaron el efecto retroactivo que tienen los exámenes de competencias mínimas exigidos para graduarse de enseñanza media. Encontraron que el efecto era mayor en las cohortes de edad más recientes, probablemente porque los incentivos han crecido en la medida en que han empezado a operar realmente las consecuencias. Los alumnos de 4° básico que estudian en estados con ECM aventajan en alrededor de un tercio de año, en todas las asignaturas, a aquellos que no tienen ECM. En 8° básico la ventaja adquirida por los alumnos sometidos a ECM es de un cuarto de año en las asignaturas de matemática y ciencias, y en el caso de lectura es de un

³⁷ L. F. Winfield, "School Competency Testing Reforms and Student Achievement: Exploring a National Perspective" (*Educational Evaluation and Policy Analysis*, 12, 1990), citado por S. F. Marion y A. Sheinker. *Issues and Consequences for State Level Minimum Competency Testing Programs* (1999).

octavo de año. En este estudio se controla por índices de educación de los padres, proporción de niños en pobreza, porcentaje de alumnos afroamericanos, hispánicos y asiáticos.

C. Estudios longitudinales con muestras nacionales en EE.UU.

Para estudiar el efecto de los ECM exigidos para graduarse de enseñanza media, Bishop, Mane y otros (2001) hacen un análisis longitudinal estableciendo las ganancias de aprendizaje entre 8° y 12° año por alumno, lo que implica que se controla por la calidad de la educación previa del alumno. Además toman en cuenta una serie de variables de los alumnos, familia, escuela y estado. Un aspecto interesante de este estudio es que incluye en las evaluaciones tanto a los alumnos que estaban terminando enseñanza media como a los que habían desertado, con el fin de controlar parte de los efectos negativos que pueden producir este tipo de exámenes. Utilizan los datos del NELS³⁸ de 1988-1992, en las cuatro asignaturas básicas de enseñanza media. Observan que los ECM tienen un efecto positivo pero no significativo cuando se promedian las cuatro asignaturas. Los alumnos que tenían rendimientos iniciales promedio, con ECM alcanzaron ganancias positivas y significativas en matemática y ciencias sociales.

B. Jacob (2000) utiliza la misma serie de datos que Bishop, Mane y otros (2001), controla por variables similares e incluye correcciones por sesgos de selección y sesgo de error de medición, y no observa diferencias significativas entre los alumnos con ECM y los sin ECM. Este estudio, sin embargo, se realizó durante el período inicial de aplicación de los ECM y se centró en una sola generación. Hay que tener presente que en ese período los estándares eran muy bajos en gran parte de los EE. UU. y, por otro lado, que los ECM aún no tenían consecuencias para los alumnos. También hay que tener cautela con la interpretación de estos datos, dada la evidencia antes expuesta de que los exámenes externos aplicados al final de la enseñanza media afectan al comportamiento de los alumnos desde antes de 8° básico, por lo tanto los resultados pueden estar sesgados contra los ECM.

Usando los datos del NELS-88, Jacob (2000) constata que los exámenes de competencias mínimas tienen efectos en los salarios que recibirán después los estudiantes y que esos salarios han aumentado desde 1981. En las primeras mediciones no se encontró un efecto en los salarios, sin embargo en la medición de 1993 los ex alumnos de los estados con ECM que tenían un rendimiento inicial promedio ganaban un 7% más por mes y

³⁸ El NELS es el National Educational Longitudinal Survey.

anualmente un 9% más. Bishop y Mane y otros (2001) piensan que este incremento se debe a que los empleadores se demoraron en adoptar políticas salariales que reconocieran las diferencias entre los egresados que habían superado la valla de los ECM y aquellos que no lo habían logrado. Ello se debió, en parte, a que tomó tiempo difundir la existencia del sistema entre el empresariado y en parte a que en el sector productivo los trabajos empezaron a ser cognitivamente más complejos y demandantes.

En síntesis, si bien los estudios son preliminares, se puede observar en la mayoría de ellos un efecto de signo positivo de los ECM, aunque de distinta magnitud entre ellos. En general el efecto de los ECM es mayor en los cursos de enseñanza básica que a la salida de enseñanza media, y los alumnos de habilidades bajo el punto de corte de la prueba son quienes aparecen beneficiados más consistentemente en las distintas investigaciones. También los estudios permiten pensar que hay un efecto diferencial según el tipo de consecuencias establecidas. A continuación veremos los estudios sobre los exámenes externos de finales de curso. Ellos están afectados a las mismas consideraciones anteriores, sin embargo muestran efectos mayores.

3.2. Los exámenes externos finales basados en el currículo (EEFBC)³⁹

La mayoría de las investigaciones sobre EEFBC han sido realizadas por John H. Bishop y su equipo, ya que son ellos los que han diferenciado los exámenes externos finales basados en el currículo de los exámenes de competencias mínimas, y los que han dedicado parte importante de sus investigaciones al tema de los efectos de los sistemas de EEFBC. Bishop⁴⁰ los define como:

1. Colecciones de exámenes que señalan los logros de aprendizaje de los alumnos y asocian consecuencias reales para los estudiantes. Un ejemplo son los Regents Exams de Nueva York⁴¹.

³⁹ En inglés el término empleado es Curriculum-Based External Exit Examinations (CBEEE).

⁴⁰ Véanse J. H. Bishop, "The Effect of Curriculum-Based External Exit Exam Systems on Student Achievements" (1998), y J. H. Bishop y otros, "Diplomas for Learning, Not for Seat Time: The Impacts of New York Regents Examinations" (2000b).

⁴¹ Los Regents Exams se toman a lo largo de la educación media. Un alumno típico, orientado a la universidad (*college*), tomaría el examen de matemática y de ciencias de la

2. Definen el logro de acuerdo a estándares externos no relativos a los estudiantes de la clase ni del colegio. De este modo se pueden hacer comparaciones entre colegios y alumnos de distintos establecimientos sobre una base de datos común. La mayoría de las veces se refieren a pruebas de criterio pero algunas veces se utilizan pruebas referidas a normas estables en el tiempo.

3. Se organizan por disciplinas y se asocian a contenidos y temas específicos como también a una secuencia definida de cursos. De este modo la responsabilidad de preparar a los alumnos recae en un grupo de profesores reducido y claramente identificables. Cuando los alumnos pueden tomar cursos electivos, se asocian exámenes específicos. La alineación entre instrucción y evaluación se maximiza y la responsabilidad se intensifica.

4. Los EEFBC señalan múltiples niveles de rendimiento. Éste es un punto importante, porque cuando sólo hay la posibilidad de aprobar/reprobar, como en los exámenes de competencias mínimas, la tendencia es bajar el estándar, de manera que una mayoría pueda pasarlo. En cambio los EEFBC miden y señalan el intervalo total de rendimiento escolar de cada asignatura, por lo tanto pueden abarcar preguntas más difíciles y temas cognitivamente más demandantes. De este modo, todos los alumnos tienen incentivos para esforzarse más, no sólo los que están bajo el mínimo esperado sino que también los que están en el tope de la distribución.

5. Un sistema de exámenes externos finales basados en el currículo evalúa a todos los alumnos, aunque no necesariamente con la misma prueba. Muchos países ofrecen distintos niveles de dificultad para la misma asignatura y permiten la elección de ramos y temas. Los exámenes de ingreso a los colegios secundarios de elite, las pruebas de ingreso a la universidad y la certificación de cursos avanzados no se consideran EEFBC, porque tienen influencia sólo en los estándares de los alumnos que están en el tope de la distribución.

Ejemplos de estas pruebas los encontramos en Francia con el Baccalauréat y el Brevet; en Alemania con el Abitur y en Israel con el Bagrut; en Inglaterra y Gales con el Certificado General de Enseñanza Secundaria (GCSE) y los A-Levels; en Estados Unidos están los Regents Exams en

Tierra en el 9º año, matemática, biología y estudios globales en el 10º año; matemática, química, inglés, historia americana y lengua extranjera en el 11º año, y física en el 12º año. Un 89% de los alumnos de Nueva York toma los exámenes de matemática y alrededor de un 60% los de historia, biología e inglés. Los alumnos que deciden no examinarse con los Regents Exams tienen que rendir los Regents Competency Tests, que se ajustan al esquema de exámenes de competencias mínimas. J. H. Bishop y F. Mane, "The Impact of Minimum Competency Exam Graduation Requirements on College Attendance and Early Labour Market Success of Minority Students" (1998).

CUADRO N° 1: CLASIFICACIÓN DE PAÍSES, ESTADOS Y PROVINCIAS SEGÚN PRESENCIA DE EEFBC* EN MATEMÁTICA Y CIENCIAS

Presencia de EEFBC en ciencias y matemática	Presencia de EEFBC en matemática	Presencia de EEFBC sólo en algunas provincias o estados del país	Muestra de países sin EEFBC
Bulgaria República Checa Hungría Lituania Eslovenia República Eslovaca Rusia	Rumania		Latvia
Dinamarca Países Bajos Holanda Inglaterra Irlanda Escocia	Noruega Islandia Francia	Australia Alemania Suiza	Bélgica Portugal España Suecia** Grecia Chipre
Nueva Zelanda			Filipinas
Hong Kong Japón Corea Singapur Tailandia Malasia			
Irán Israel			
Colombia		Estados Unidos*** (Texas, Carolina del Norte; Nueva York). Canadá (Alberta, British Columbia, Manitoba, Newfoundland, Quebec y colegios francófonos de New Brunswick).	Chile

* Los exámenes de admisión a las universidades no se consideran exámenes externos finales basados en el currículo (EEFBC) a menos que se utilicen al mismo tiempo para certificar la enseñanza media. Para clasificar como EEFBC un examen es requisito que lo rindan todos los alumnos.

** Suecia es un caso bastante particular, no aplica exámenes externos pero entrega pautas de evaluación que marcan el estándar exigido para cada curso para evitar la inflación de notas. El National Board of Education elaboró un examen de alternativas y de respuestas cortas para “ayudar a los profesores a evaluar apropiadamente”, el cual es obligatorio y condiciona las notas finales del curso, ya que éstas no deben desviarse en más de 0,2 puntos del promedio estandarizado obtenido en las pruebas.

*** En Estados Unidos en el año 2001, 18 estados retuvieron los diplomas de graduación dependiendo de exámenes estatales de competencias mínimas, tres estados basaron en EEFBC la promoción al curso siguiente en enseñanza básica y sólo tres estados tenían EEFBC. En un futuro tendrán exámenes de EEFBC: Maryland, Mississippi, Oklahoma, Arkansas, Tennessee y Virginia.

Fuente: Elaboración propia a partir de NGA Center for Best Practice, “Graduation Exit Exam Matrix”; Britton y Raizen, *Examining the Examinations* (1996), y Bishop, “The Impact of Curriculum-Based External Examinations on School Priorities and Student Learning de Bishop” (1996).

Nueva York y los Golden States de California; y los EOC de Carolina del Norte; y en Canadá son varias las provincias que los tienen. El Cuadro N° 1 hace un listado de países que tienen EEFBC al menos en ciencias y matemática.

Evidencia empírica de las ventajas de los exámenes externos finales basados en el currículo (EEFBC)

A. *Estudios internacionales*

En los últimos 10 años, John H. Bishop junto a otros investigadores ha realizado múltiples estudios que demuestran que los países y distritos que tienen exámenes externos finales basados en el currículo tienen efectos positivos en el rendimiento de los alumnos. Utilizando los resultados de rendimiento en el TIMSS⁴² de 1994-1995 de 8^{os.} años de enseñanza básica como variable dependiente, compara países que aplican EEFBC con los que no lo hacen, controlando por ingreso per cápita ajustado por poder de compra y por la presencia de cultura asiática. El impacto de los EEFBC es equivalente al avance que se logra en 1,3 año de enseñanza en EE.UU. en ciencias y 1 año en matemática⁴³. Esta medida toma como referencia cuántos puntos aumentó EE.UU. entre el 7° y 8° grado en la medición del TIMSS, de modo que esa unidad determina un año.

Wößmann (2001) repite el estudio utilizando la misma serie de resultados del TIMSS, más los datos obtenidos en los cuestionarios que este estudio aplica a establecimientos, profesores y alumnos, más los indicadores de países construidos por la OECD. De este modo puede controlar por características de los alumnos y de la familia, por recursos pedagógicos (número de alumnos por clase, número de horas de clases, materiales didácticos, etc.), características de los profesores, por el tipo de institucionalidad educativa de cada país (nivel de responsabilidad de administración de los recursos y gestión de los colegios, presencia de exámenes externos centralizados). En el caso de esta investigación, el total de variables analizadas explica un 60% de la varianza. Tomando en cuenta todas estas variables, el estudio concluye que en los países en que hay EEFBC se observa un impacto positivo y significativo en el rendimiento. Los estudiantes con exámenes

⁴² El Tercer Estudio Internacional de Matemática y Ciencias (TIMSS) es un estudio que compara el rendimiento en treinta y ocho países en 3°, 4°, 7°, 8° y 12° año de enseñanza. Para una referencia del TIMSS véase, H. Beyer, "Falencias Institucionales en Educación: Reflexiones de Propósito de los Resultados del TIMSS" (2001).

⁴³ Véase J. H. Bishop, "The Effect of National Standards and Curriculum-Based Exams on Achievement" (1997).

centralizados tienen una puntuación de 16,1 puntos superior en matemática (alrededor de 1/3 de desviación estándar) y 10,7 puntos en ciencias (1/5 de desviación estándar). Quizás este estudio controla por demasiadas variables y algunas de ellas no son independientes del todo de la instauración de los exámenes externos, por lo tanto podrían sesgar ligeramente los resultados en contra de ellos.

En otro estudio realizado por Bishop (1996), usando los datos del International Assessment of Educational Progress (IAEP) de 1991, con la población de 13 años, en 15 naciones se encontró que los alumnos que seguían estudios en sistemas con exámenes externos superaban a los alumnos que no los tenían en 2 grados escolares de EE.UU. en matemática y dos tercios de grado en ciencias y geografía. Estas diferencias son grandes y significativas. En este estudio se controló por producto geográfico bruto per cápita corregido por poder de compra y cultura asiática.

B. Estudios en EE. UU.

Bishop, Mane y otros (2001) analizan dos series de datos, los del NAEP de 4° y 8° básico de 1992, 1994, 1996, 1998 en lectura, matemática y ciencias, y los datos del estudio longitudinal NELS-88 que estima las ganancias logradas desde 8° al 12° año por alumno. Estas bases de datos recogen además una serie de variables de los alumnos, padres, familia, escuela e institucionales. Bishop y su equipo encuentran que el híbrido constituido por exámenes de competencias mínimas más exámenes externos finales basados en el currículo tuvo un efecto grande y significativo en el rendimiento en comprensión de lectura, matemática y ciencias en 1994, 1996 y 1998. Los efectos fueron un avance del orden de medio a tres cuartos de año escolar. Los autores comprueban, además, que los efectos crecen con el tiempo. En Carolina del Norte las pruebas se instauraron entre 1988 y 1991, observándose menores efectos en 1992, recién a un año de implementada la política; en cambio en 1998 se observa un efecto mayor. Esto es razonable, ya que los alumnos y profesores requieren readecuar sus esfuerzos desde cursos tempranos para enfrentar con éxito los exámenes finales.

C. Estudios en Canadá

Bishop (1996) realizó una investigación en las diferentes provincias de Canadá con el fin de estimar el impacto que tienen los exámenes externos de fin de la enseñanza media en el rendimiento de los alumnos de 8°

básico. Es decir, intentó analizar el efecto retroactivo de los EEFBC. Para esto utilizó los datos de IAEP (International Assessment of Educational Progress) recopilados durante 1990-1991. Esta investigación pudo distinguir tres tipos de provincias: a) las que tenían exámenes que implicaban costos para los alumnos (por ejemplo, la nota del examen se contaba como un 50% de la nota final en los estados de Quebec, Alberta y Newfoundland y un 40% en British Columbia); b) las que tenían exámenes pero los resultados no se reportaban a los alumnos ni se consideraban en las notas finales de los cursos, y c) las que no tenían exámenes externos. Se controló por variables demográficas de los alumnos y nivel educacional de los padres. El nivel de gasto en educación de las distintas provincias no se controló porque en Canadá es relativamente parejo. Se comprobó que la presencia de exámenes externos en 12° grado afectaba significativamente al rendimiento de los alumnos de 8° año básico. Las provincias con EEFBC tienen una diferencia a su favor de 24,3% de una desviación estándar, esto es, cerca de 4/5 de un año escolar de EE.UU. en matemática y 17,6% de desviación estándar en ciencias, lo que equivale a 2/3 de grado de EE.UU. Junto con el análisis de la influencia que tienen los exámenes en el rendimiento se estudió su efecto sobre el comportamiento de los padres, alumnos y escuelas. Se constató que las provincias con EEFBC tendían a que:

- los padres conversaran más con sus hijos acerca de las clases de matemática y ciencias;
- los niños reportaran, más frecuentemente, que sus padres les exigían que lo hicieran bien en matemática;
- los colegios programaran más horas de matemática y ciencias, asignaran más tareas, tuvieran mejores laboratorios y contrataran profesores especialistas en matemática y ciencias.

Estas observaciones realizadas en Canadá concuerdan con el hecho de que en los otros países que aplican EEFBC también se constata que tienden a contratar profesores especializados por asignatura en el segundo ciclo de enseñanza básica, tienen estándares más altos para entrar a la profesión docente y tienden a pagarles más a sus profesores. Sin embargo, se observa que los profesores están menos satisfechos con su trabajo, posiblemente por la presión de rendir en este tipo de exámenes.

Se podría pensar que la contratación de profesores especializados y los mayores sueldos explican las diferencias de rendimiento, sin embargo Bishop (2000b) defiende la idea de que estas diferencias son producto de la

CUADRO N° 2. RESUMEN DE LA EVIDENCIA EMPÍRICA DE LAS VENTAJAS DE LOS EXÁMENES EXTERNOS FINALES BASADOS EN EL CURRÍCULO

Autor	Muestra	Controles	Resultados*
Bishop (1997) Efecto de los EEFBC de media en rendimiento 8° básico.	TIMSS 1995. Alumnos de 13 años. 39 países y provincias.	Regresión con PGB per cápita deflactado por ppp, Dummy por país de cultura asiática.	Diferencias de 1 año de estudio en matemática y de 1,2 en ciencias.
Bishop (1996) Efecto de los EEFBC de media en rendimiento 8° básico.	IAEP 1991. Alumnos de 13 años. 15 países.	Regresión con PGB per cápita deflactado por ppp, Dummy por país de cultura asiática.	Ventaja de 2 años escolares de EE.UU. En ciencias se encontró un efecto no significativo.
Wößmann (2001) Efecto de los EEFBC de media en rendimiento 8° básico.	TIMSS 1995. Alumnos de 13 años (39 países). Más indicadores de la OECD.	Regresión con variables del alumno, de la familia, de la escuela, de los profesores y de la institucionalidad educativa.	1/3 de desviación en la distribución de rendimiento en el TIMMS en matemática y 1/5 en ciencias.
Bishop, Mane y otros (2001) Efecto de los EEFBC de media en rendimiento de enseñanza media.	NAEP 92 al 98 y NELS-88. Alumnos de 17 años y seguimiento de alumnos de 8° a 12°.	Regresión con variables del alumno, de la familia, de la escuela, de los profesores y de la institucionalidad educativa.	En ciencias, lectura y matemática se observa un avance de medio a tres cuartos de año escolar de EE.UU.
Bishop (1996) Efecto de los EEFBC de media en rendimiento 8° básico.	Progress IAEP 90-91. Alumnos de 13 años de distintas provincias de Canadá.	Regresión con variables demográficas de los alumnos, educación de los padres. El nivel de gasto en educación no se controló porque es relativamente constante en las provincias de Canadá.	Ventaja de 4/5 de año escolares de EE.UU. en matemática y en ciencias 2/3 de año.

* Los resultados se expresan en términos del avance en años de estudio que implica un sistema u otro. El cálculo se realiza sobre la base del promedio de puntos que avanzan los alumnos en un año en Estados Unidos en referencia a la prueba utilizada.

influencia de los exámenes externos⁴⁴. Sostiene que cuando los padres toman conciencia de la importancia de la educación, porque enfrentan el hecho de que el no cumplir con las exigencias impuestas a los niños tiene consecuencias, empiezan a presionar al sistema para que invierta en mejo-

⁴⁴ Véase la discusión al respecto en D. Ravitch (ed.), "Beep Conference on National Students" (2000) y J. H. Bishop, "The Impact of Curriculum-Based External Examinations on School Priorities and Student Learning" (1996), sección "New York State vs. The Rest of The United States". Bishop sostiene que las diferencias de rendimiento no se podrían atribuir directamente al hecho de que unos estados tienen más recursos para invertir en educación que otros. Cree más bien que son las prioridades que establecen los EEFBC las que estarían afectando a los resultados. Por ejemplo, Nueva York, estado que aplica EEFBC, paga más a sus maestros de educación media que otros estados más ricos como Massachusetts, New Jersey y New Hampshire, que no tienen EEFBC. En esta comparación toma en cuenta las diferencias del costo de la vida entre los estados.

res profesores y para que se destinen los recursos hacia lo que reporta beneficios académicos más que hacia otros ítemes.

Por último es importante señalar que los exámenes de competencias mínimas y los EEFBC aumentan la probabilidad de asistir y mantenerse en la universidad. Esto parece lógico: si los alumnos tienen que esforzarse más durante la enseñanza media seguramente quedan mejor preparados para enfrentar las exigencias de las pruebas de selección y las exigencias de las carreras. Llama positivamente la atención que los alumnos con rendimientos bajos al ingresar a enseñanza media en los estados con ECM y EEFBC tienen una mayor probabilidad de estar en la universidad que los alumnos de rendimientos bajos de estados sin este tipo de exámenes⁴⁵. Si se analiza la asistencia a la universidad dos años después de la graduación de la generación, los alumnos de “C-” al ingresar a educación media tienen una probabilidad mayor de estar en la universidad que los alumnos sin exámenes externos estimada en 2,3 puntos porcentuales. Esta probabilidad es aun mayor para estos alumnos en el estado de Nueva York, que tiene el híbrido ECM y EEFBC, que se estimó en 6 puntos porcentuales. La probabilidad para los otros rangos de habilidad también es más alta en los estados que ponen exigencias mínimas para graduarse, para los alumnos con “B” era de 4,4 puntos y con “A” de 3,3 puntos.

3.3. Contribución de los ECM y EEFBC en relación con otras medidas de rendición de cuentas

Bishop, Mane, Bishop y Moriarty (2001) estudian el impacto de distintas medidas de rendición de cuentas en el rendimiento de los alumnos en la evaluación nacional del NAEP de 1996 en matemática y ciencias y en la de 1998 de comprensión de lectura en 4° y 8° básico. Las medidas que revisan son la publicación de los resultados de las escuelas, premios a los colegios que tienen buenos desempeños, sanciones a los establecimientos que no alcanzan las metas, asistencia a los colegios que no tienen buenos resultados y la existencia de ECM. Lo primero que concluyen es que estas variables tienden a estar asociadas⁴⁶: los estados tienden a

⁴⁵ Se hace una regresión logística utilizando como variable dependiente la ganancia de aprendizaje entre 8° y 12° en el National Educational Longitudinal Survey. El estudio analizó diferenciadamente los efectos de los ECM según el rendimiento inicial de los alumnos en enseñanza media (alumnos con A, B, C y C menos). Véase Bishop y Mane (1998).

⁴⁶ Los ECM tienen una correlación de 0,526 con sanciones + premios; 0,544 con sanciones; 0,3 con la publicación de reportes de resultados. La asistencia a los colegios que fallan tiene una correlación 0,575 con la publicación de reportes de resultados; 0,709 con sanciones a colegios que fallan.

aplicar estas medidas en paquete. Como dijimos, esto es razonable, ya que aplicar consecuencias sólo a los alumnos o sólo a los profesores/escuelas resultaría injusto porque el desempeño de cada uno depende del otro. Por lo tanto, teóricamente se debieran aplicar medidas conjuntas que exijan cuentas tanto a profesores como a alumnos. Aún así, este grupo de autores hace una regresión para estudiar qué parte de la varianza explica cada una. Para hacer estimaciones no sesgadas del efecto particular de cada una de las medidas de rendición de cuentas se analizaron estados en que se hallaba presente una de las medidas en ausencia de las otras, utilizando las grandes muestras de datos aportadas por el NAEP en diferentes años y asignaturas. De esta manera tratan de resolver el problema de la alta asociación de las variables. La política que claramente tiene los efectos más grandes es el híbrido de ECM/EEFBC de Nueva York y Carolina del Norte (tienen un avance de 45% de un año escolar en EE. UU. en matemática y ciencias y 65% en lectura respecto a los estados demográficamente similares que no aplican la política). El efecto retroactivo de los exámenes de competencias mínimas de enseñanza media en el rendimiento de básica tiende a diluirse cuando se controla por otras medidas de rendición de cuentas. En este estudio, la segunda política más efectiva es imponer premios en conjunto con sanciones a los profesores y colegios. Publicar los resultados de las escuelas y de los niños es una medida necesaria para aplicar las dos políticas anteriores, pero por sí misma no tiene efectos significativos en el rendimiento de los alumnos⁴⁷. Sin embargo, los resultados de esta investigación deben ser interpretados con cautela porque se está midiendo el efecto diferido de los exámenes finales de enseñanza media en 8° básico y, en cambio, las otras medidas de rendición de cuentas se aplican directamente en el ciclo básico.

Sin duda las investigaciones no son ideales, sin embargo la evidencia y la experiencia comparada apuntan al rol positivo que juegan los exámenes externos con consecuencias individuales en el mejoramiento de la calidad de la educación.

⁴⁷ Este resultado no es estrictamente extrapolable a Chile, ya que en Estados Unidos los padres tienen menos libertad para elegir los colegios de sus niños porque deben matricularse en el que queda más cerca de su hogar. Por lo tanto, la información no la pueden utilizar para retirar a sus hijos de aquellas escuelas que lo hacen mal. En cambio en Chile los padres tienen la libertad de escoger y cambiarse de colegio libremente, llevando el subsidio que le corresponde a su hijo a la escuela que le parezca más conveniente. Posiblemente, entonces, mayores grados de información pueden tener un efecto mayor en nuestro país.

IV. CRÍTICAS FRECUENTES Y POSIBLES DESVENTAJAS DE LOS ECM Y EEFBC

A continuación se presenta un listado de las críticas y contraargumentos generados en el intenso debate surgido en Estados Unidos a raíz de la imposición de exámenes externos con consecuencias para los alumnos. Se critican los exámenes externos porque:

1. *Reducen el aprendizaje a conceptos de orden simple, a la memorización sin comprensión.* Ésta es una crítica clásica que se les hace a las pruebas de alternativas múltiples que frecuentemente se aplican en este tipo de exámenes masivos. Esta inquietud no es totalmente válida porque estas pruebas pueden medir una amplia gama de destrezas, conocimientos y niveles de pensamiento si están bien planteadas. Un ejemplo de preguntas bien planteadas se encuentra en la prueba TIMSS.

Por otra parte, la mayoría de los países europeos y exámenes de calidad como los del bachillerato internacional utilizan el formato de ensayo en sus evaluaciones. En Estados Unidos, los estados están empezando a complementar las pruebas de alternativas múltiples con otro tipo de formatos que incluyen preguntas abiertas de respuesta corta o extendida, y con pruebas de desempeño en las cuales los alumnos tienen que conducir experimentos, disertar, argumentar. También se ha incorporado el uso de portafolios para evaluar trayectorias de aprendizaje en redacción y arte. En los últimos 15 años el desarrollo del tema de evaluación ha sido explosivo dada la importancia que ha adquirido la rendición de cuentas y por consiguiente la calidad técnica de las evaluaciones está mejorando considerablemente⁴⁸.

Si bien el formato de las preguntas incide en la metodología de enseñanza, lo que más puede afectarlas es la calidad de las preguntas. Preguntas irrelevantes, capciosas, muy puntuales, idiosincráticas, inducirán a los estudiantes y alumnos a adivinar los trucos y a memorizar “trivia” más que aprendizajes útiles y profundos. Validar las preguntas internacionalmente y con expertos de variadas orientaciones es fundamental para asegurar que la educación impartida en los establecimientos se oriente hacia lo relevante.

⁴⁸ Un análisis interesante sobre tendencias en el tipo de diseño de pruebas de evaluación externa se encuentra en R. J. Marzano, *Models of Standards Implementation: Implications for the Classroom* (1998). Para un completo análisis de las críticas más frecuentes a las pruebas de evaluación estandarizadas, véase R. P. Phelps, “Why Testing Experts Hate Testing” (2002).

Bishop ejemplifica la idea de que los exámenes finales externos basados en el currículo pueden promover buenas prácticas educativas y señala que, en las encuestas del estudio TIMSS, los alumnos de países con exámenes externos basan su estudio en menor grado en la memorización que los de países sin esos exámenes.

2. *Reducen la educación sólo a lo que será medido.* La visibilidad de los resultados de los exámenes de rendición de cuentas y las consecuencias asociadas a ellos generan incentivos poderosos para concentrar la enseñanza sólo en las asignaturas evaluadas. La evidencia muestra que esto es así; Deere y Strayer (2001) observan cómo en el estado de Texas mejora más el rendimiento en aquellas asignaturas evaluadas que se tomaban en consideración para el sistema de premios y sanciones que en aquellas asignaturas que eran evaluadas pero no se utilizaban para la ponderación. De esto se concluye, en primer lugar, que los sistemas de rendición de cuentas deben considerar todos aquellos ramos que se juzgan primordiales de mejorar. Los sistemas que han implementado este tipo de exámenes hace más tiempo evalúan un número más extenso de asignaturas. Sin embargo muchos de los sistemas que recién están iniciándose priorizan aquellos ramos que consideran instrumentales y más urgentes de modificar. La mayoría comienza por matemática, lectura y redacción. Los sistemas de rendición de cuentas asignan prioridad a lo académico, ya que es uno de los aspectos más difíciles de evaluar por los padres. En relación con las áreas deportivas y artísticas, normalmente se utilizan otros mecanismos para evaluarlas y realzar su importancia. En los países con mayor trayectoria en evaluación, se hace por la vía de presentaciones, exposiciones, concursos y competencias. Estas actividades hacen más transparentes ante los padres la gestión de los establecimientos en esas áreas y se pueden considerar una forma de rendición de cuentas. Del mismo modo, los aspectos formativos, los disciplinarios, de infraestructura y las actividades extracurriculares resultan accesibles y evaluables para los padres.

3. *Hay aumento espurio del aprendizaje.* Detrás de esta crítica está la noción de que una enseñanza dirigida hacia una prueba determinada (*teaching to test*) puede producir un aumento circunscrito del aprendizaje por familiaridad con el instrumento. Por otra parte, el hecho de que los incentivos para hacer trampa aumentan cuando las consecuencias son más fuertes; también contribuye a la aprensión de que los resultados puedan no reflejar aprendizajes reales. Un buen diseño del sistema de evaluación debería hacerse cargo de estos problemas.

En primer lugar, los exámenes tienen que considerar que uno de los principios que están detrás del movimiento de los estándares es que los profesores deben orientar la enseñanza hacia las pruebas⁴⁹. Por lo tanto, la familiaridad con lo que se evaluará es deseable y los contenidos y el tipo de preguntas no deben ser un secreto. Publicar las evaluaciones anteriores es una práctica difundida entre los países que aplican exámenes externos. Sin embargo, esto implica crear nuevos ítemes equivalentes cada vez y la seguridad de que el formato y el conjunto de tareas exigidas apuntan a las destrezas y conocimientos relevantes.

En segundo lugar, la aplicación de pruebas paralelas muestrales ayuda a detectar si las evaluaciones individuales están reflejando un aprendizaje circunscrito al tipo de preguntas realizadas o uno más generalizado. La participación en pruebas internacionales es una buena práctica en este sentido.

4. *Restringe la libertad de enseñanza.* Ésta es una de las aprensiones centrales y más frecuentes que surgen al momento de instaurar exámenes externos. Las evaluaciones que tienen consecuencias importantes para los alumnos y los establecimientos condicionan lo que se enseña. Chile tiene consagrado por un lado el derecho a la libertad de enseñanza y por otro lado garantiza el derecho a la educación escolar gratuita y pública. Por supuesto, hay una tensión entre estos dos puntos, la cual se intenta resolver con un currículo de objetivos mínimos y contenidos fundamentales. Sin embargo, en la práctica nuestro currículo, es amplio y deja pocos márgenes de libertad. La válvula de escape a este *impasse* ha sido la posibilidad de interpretar más libremente el currículo, dada la ausencia de un sistema de rendición de cuentas eficiente. Sin embargo, si se aplican exámenes con consecuencias para los alumnos es necesario hacerse cargo del punto.

En primer lugar, lograr el estándar exigido en las evaluaciones externas no debiera cubrir todo el horario disponible en cada asignatura, es necesario asegurar una proporción de tiempo para que cada establecimiento ponga los acentos donde crea que es importante. En este sentido, si en Chile se quisiera implementar un sistema de exámenes externos con consecuencias individuales, sería necesario revisar el currículo y definir sobre esa base los estándares que realmente se consideran elementales de lograr. Un método complementario utilizado para asegurar márgenes de libertad consiste en evaluar sólo destrezas generales, respecto a las cuales hay amplios grados de acuerdo acerca de su importancia y el modo de evaluarlas. En

⁴⁹ Véase la discusión sobre el “teaching to test” de Wiggins, en R. J. Marzano, “Models of Standards Implementation: Implications for the Classroom” (1998), p. 14.

matemática, comprensión de lectura y redacción se han logrado altos grados de consenso en la mayoría de los países con buenos niveles educacionales. Para abordar temas más específicos de cada asignatura, en general los países europeos y asiáticos optan por permitir que los alumnos escojan entre distintos temas y áreas, como también entre distintos niveles de dificultad. Por ejemplo, en Singapur los alumnos del 10° grado tienen que rendir exámenes elaborados por el Cambridge Syndicate de Inglaterra. Escogen entre los “GCE Normal Level” (académicos o técnicos) y los “GCE Ordinary Level”, que son más difíciles. En cada una de estas alternativas hay a su vez posibilidades de elección. Por ejemplo, en historia hay una parte general para todos sobre Malasia-Singapur (1819-1963), en la cual se acotan 6 temas, y una parte en la que los alumnos son examinados en el tema de un país del sudeste asiático que cada uno escoge previamente para estudiar en profundidad⁵⁰.

Otra manera de resguardar la libertad educacional de las escuelas es permitir la elección entre diferentes organismos examinadores. Por ejemplo, Inglaterra tiene distintos organismos examinadores que están debidamente acreditados; Virginia permite reemplazar su examen final de enseñanza media por equivalentes reconocidos, como el International Baccalaureate, las pruebas del ETS de Advanced Placement⁵¹ y el SAT-II. Cuando se permite escoger entre exámenes es importante conjugar la libertad con el garantizar la equivalencia en calidad entre las distintas evaluaciones. También es crucial resguardar que la información acerca del nivel de logro alcanzado por los alumnos se mantenga relativamente simple, de modo que el mercado laboral pueda interpretarla; si hay gran variedad de exámenes sin jerarquía los empleadores no sabrán qué valor asignarles a determinados puntajes.

Para permitir la experimentación, algunos países liberan a las escuelas que han alcanzado buenos rendimientos de la exigencia de rendir exámenes externos en algunos tramos específicos.

5. *La presión por el logro interfiere con el desarrollo de la motivación intrínseca.* La motivación intrínseca por el aprendizaje es el gusto por

⁵⁰ Para una descripción de los exámenes rendidos en Singapur se puede consultar en la página web del Ministerio de Educación de Singapur. www.moe.edu.sg/exams/2002.

⁵¹ Según datos del ETS (Educational Testing Service), el organismo que administra la prueba del Advanced Placement, se rinden 1.400.000 exámenes anuales. El reconocimiento de estas pruebas como un instrumento útil para exigir cuenta a los profesores de cursos específicos y a los alumnos, las está haciendo cada vez más populares, lo cual hace probable que un número cada vez mayor de estados las acepte como un equivalente de sus propios exámenes.

el conocimiento mismo y el interés por el tema propiamente tal, básicamente la curiosidad. Generar estas actitudes es uno de los ejes de la educación, y por lo tanto es fundamental cuidar este aspecto. Es cierto que exigencias desmedidas con costos altos asociados al riesgo de no alcanzarlas pueden producir grandes dosis de ansiedad, lo que interfiere con el gusto por aprender. Sin embargo, si las exigencias son graduadas y las materias se ofrecen con una buena didáctica, no tendría por qué ser de ese modo. Bishop (1996) comprueba que los alumnos que viven en estados o países con exámenes externos, controlando por ingreso per cápita y otras variables demográficas, leen más frecuentemente por placer y eligen para ver series televisivas del tipo del National Geographic. Al parecer, los niños que más saben quieren aprender más.

6. *Suscita el aumento de la deserción y la repitencia.* Es un hecho que la deserción y la repitencia aumentan entre los alumnos que tienen un rendimiento bajo inicial cuando se les impone exigencias de logro. Lillard y de Cicca (2000); Bishop y otros (2000a); Jacob (2000) estudian cómo afectan en la deserción los ECM y los EEFBC aplicados al final de la enseñanza media. Controlando por distintas variables sociodemográficas, todos concluyen que los ECM no inciden en la deserción a excepción de los alumnos que tenían un rendimiento bajo en la enseñanza básica. Bishop y su equipo estiman que en los estados con ECM hay una probabilidad de que un 31% de los alumnos que tienen rendimiento muy bajo (C-), al ingresar a enseñanza media no completen sus estudios. En cambio, esta proporción es más baja en los estados sin ECM, ya que se estima que el 24% de estos alumnos no terminaría sus estudios escolares.

También Bishop y otros (2000a) analizaron el tiempo que les toma a los alumnos graduarse en los estados con ECM y sin ECM. Los alumnos con promedios iniciales “B”, “C”, y “C-” se demoran más en graduarse cuando deben rendir con éxito una prueba de contenidos mínimos. En promedio, en los estados con ECM un 28% de los alumnos destina más años que lo regular para obtener su certificado de graduación versus un 17% en los estados sin esta exigencia⁵². En los estados con ECM hay un ligero

⁵² Los sistemas educacionales que aplican exámenes externos con consecuencias para los alumnos se asimilan al concepto de “aprendizaje para el dominio”, según el cual los alumnos no pueden pasar a las etapas superiores mientras no hayan alcanzado los conocimientos que se consideran prerrequisitos. Pedagógicamente esto tiene grandes ventajas para el alumno, porque enfrenta tareas para las cuales está preparado. Sin embargo, los alumnos tienden a graduarse más tarde. Si se revisa la sobreedad de los alumnos matriculados en los últimos años de enseñanza media en países que tienen exámenes exigentes para graduarse, nos podremos dar cuenta de que muchos estudiantes han optado por permanecer en el sistema con el fin de lograr los objetivos que se esperan. En Francia el 34% del corte de edad tiene

aumento de los alumnos que terminan su educación sólo con un diploma de asistencia en vez del certificado de graduación. Entre los alumnos con una “C-” inicial obtienen este diploma un 10% en los estados con ECM versus un 7,8% en los estados sin ECM.

Al comparar los indicadores del estado de Nueva York, que tiene un sistema de EEFBC, con los indicadores de estados sin exámenes de competencias mínimas, se observa para el caso de Nueva York una probabilidad mayor de desertar, de obtener un diploma de asistencias en vez de certificado de graduación y un incremento de las posibilidades de graduarse más tarde. También se observó que había un aumento significativo de alumnos que se graduaban tempranamente. Sin embargo, en el estado de Nueva York una proporción muy alta de alumnos de minorías ingresan a la universidad y tienen éxito en ella.

Dada la realidad de que hay un aumento en la deserción y la repitencia, los sistemas que exigen requisitos mínimos para graduarse implementan una serie de medidas para evitar estos problemas. Por otra parte, hay que considerar que la educación adquirida por los alumnos en estados que ponen condiciones para graduarse es de más calidad por año cursado, por lo tanto los alumnos que no completan su educación o que no logran aprobar el mínimo pueden estar por sobre los alumnos de bajo rendimiento en estados sin ECM. De hecho, en los estudios realizados por Bishop y su equipo (2001), si se comparan alumnos de bajo rendimiento inicial al ingresar en enseñanza media controlando por educación de los padres y otras variables demográficas, los que habían estudiado bajo el régimen de ECM ganaban 7% mensual más inmediatamente después de la graduación, probablemente al reconocérseles mayores destrezas y hábitos de trabajo.

Entre las medidas para evitar deserción y repitencia encontramos:

- Reconocer diferenciadamente a los alumnos que desertan, de aquellos que no han podido cumplir los mínimos pero que permanecen en el sistema, y a éstos a su vez de los que han alcanzado los objetivos de logro esperados. De este modo se reconoce el esfuerzo y la constancia. Si las señales son claras, el mercado las sabrá reconocer y generará incentivos para que los alumnos completen sus estudios.

19 años en el grado 12 y el 12% tiene 20 años. En Alemania, 29% tiene 20 años al terminar el 13° grado, que es el último curso del sistema escolar. En contraposición al 6% de alumnos con 19 años y 2% en el 12° grado en Estados Unidos, país que en general no tiene un sistema de examinación externa exigente de los contenidos de enseñanza media. En Chile el 15% de los jóvenes de 19 años se encuentra en la enseñanza media regular y el 6% sigue estudiando en ese ciclo a los 20 años. (Los datos para Chile se obtuvieron de la base de datos de la CASEN 1998 y los del resto de los países de Britton y Raizen, 1996.)

- La mayoría de los sistemas tienen cuotas de excepción para aquellos alumnos con dificultades de aprendizaje reconocidas. Esto permite la integración de los alumnos al sistema escolar, beneficiándolos con el aprendizaje social y el desafío de sus pares. Esta medida se aplica especialmente en enseñanza básica cuando todavía no hay diferenciación por líneas de estudio. Estos alumnos reciben un diploma de completación de estudios.
- Los sistemas que se evalúan por el número de niños que alcanzan los mínimos incentivan la preocupación por aquellos que tienen déficits de aprendizaje. Se ha observado que las escuelas tienden a reaccionar más rápido frente a los alumnos que se quedan atrás y canalizan más recursos pedagógicos hacia ellos. Así, las brechas no siguen aumentando en años posteriores. Algo de esto se está observando en sistemas que instituyen exigencias para la promoción de curso. Por ejemplo, en Chicago los alumnos que no han aprendido lo requerido deben tomar cursos remediales, con lo cual ha disminuido significativamente el número de niños que deben repetir curso en años posteriores.
- La diversificación de las alternativas educacionales también es un camino común. Muchos países la tienen en enseñanza media y algunos en básica. Alemania diferencia desde temprana edad, en cambio Japón sólo en la media. De este modo, cada línea tiene objetivos específicos mínimos de buen nivel pero orientados a distintos tipos de inteligencias y talentos. Singapur, además, tiene explícitamente definidas líneas que difieren sólo en ritmo pero no en contenidos. Este esquema tiene la ventaja de dar oportunidades a todos exigiendo rigor en cada una de ellas. Esta solución tiene, sin embargo, sus detractores. Piensan que con la proliferación de líneas educacionales se aumenta la desintegración social. Habría que estudiar entonces qué produce más estratificación: una sociedad dividida de hecho entre los que saben más y los que saben menos o una que se decide por la especialización más temprana.
- Hay medidas menores pero no por ello menos efectivas; por ejemplo, 14 estados entregan licencia de conducir sólo a aquellos alumnos que asisten regularmente a clases. Y otras medidas muy focalizadas, como entregar un pago por asistir a clases a aquellos jóvenes de sectores desaventajados con alto riesgo de deserción⁵³.

⁵³ Véase la propuesta de H. Beyer en "Educación y Desigualdad de Ingresos: Una Nueva Mirada" (2000).

Sin embargo, ninguna de estas medidas puede solucionar el problema de aquellos jóvenes que prefieren, ante todo, maximizar su tiempo de ocio presente. Para ellos siempre estará la salida de retirarse de la escuela si los exigen un poco. Lo importante es prevenir esta actitud entre los alumnos más jóvenes enseñándoles a posponer las gratificaciones.

7. *Baja el desempeño de los alumnos con mejor rendimiento.* Uno de los peligros de los sistemas que aplican exámenes de competencias mínimas es el centrar todas sus energías en los alumnos en riesgo, descuidando a aquellos que tienen mejor rendimiento. Por otra parte, como los ECM fijan el estándar en lo más básico, no desafían a los que podrían mejorar con mayores exigencias. Para solucionar este problema, Betts y Costrell (2001) sugieren establecer criterios múltiples (al menos cuatro) que permitan distinguir distintos niveles de aprendizaje, con lo cual se reconocen todos los aumentos de esfuerzo. La publicación de los resultados por escuela se haría cargo de estas diferencias de rendimiento, aunque las sanciones se establecerían considerando las cifras referidas a los que no alcanzan los mínimos⁵⁴. Por otra parte, la propuesta (Fontaine, 2002) de entregar becas a los 30.000 alumnos sobresalientes en los exámenes externos de la educación subvencionada establecería un incentivo para los mejores alumnos que contribuiría a balancear una excesiva preocupación por los estudiantes de rendimiento moderado y bajo.

8. *Discrimina a los alumnos desaventajados socioculturalmente porque los mide a todos con la misma vara.* Es cierto que los niños de escasos recursos tienen que superar obstáculos más grandes para lograr lo mismo que los niños de sectores más acomodados con padres más educados y acceso a escuelas de mejor calidad. Sin embargo, lo que se exige son mínimos logrables, dados los recursos entregados a las escuelas. No exigirlos es en cierto modo resignarse a no garantizar una educación básica a todos.

Esta crítica también se hace en relación con la rendición de cuentas de las escuelas y profesores. En Chile, el Sistema Nacional de Evaluación del Desempeño (SNED) es un camino para hacer comparaciones más justas. Es una aproximación a la medición de valor agregado y considera entre otros factores el nivel sociocultural de los alumnos que atienden las escuelas.

9. *Se basa en el supuesto errado de que el bajo rendimiento de los alumnos se explica sólo por falta de esfuerzo.* Por supuesto que la

⁵⁴ Hay que recordar que las escuelas tienen la posibilidad de eximir a los niños con dificultades de aprendizaje más serias, por lo tanto, el sistema sólo sanciona por prácticas pedagógicas inadecuadas y no por dificultades de los niños.

visión de mejorar todos los problemas de la calidad de la educación a través de instaurar mejores incentivos para los alumnos es simplista. La escuela y los profesores también tienen que estar presionados a su vez para encontrar y aplicar los mejores métodos. Ellos son una contraparte indispensable.

Se argumenta que presionar no es suficiente cuando en el sistema no está instalado el conocimiento pedagógico de cómo enseñar mejor. Sin embargo, esto se asemeja al problema del huevo y la gallina: no hay herramientas adecuadas porque los actores no se movilizan para encontrarlas y no se les presiona porque no tienen las herramientas. Probablemente el camino es trabajar en ambos sentidos.

10. *Asocia consecuencias graves sobre la base de una medición.* En general, ninguno de estos sistemas exige jugarse el todo o nada en una sola medición. Todos implementan sus sistemas con posibilidades de repetir los exámenes. Los más cuidadosos evalúan a los alumnos con diversos instrumentos que apuntan a las mismas destrezas y conocimientos, pero los abordan de distintas formas. Por ejemplo, en Massachusetts los alumnos utilizan en conjunto pruebas de ensayo, de alternativas y de respuestas abiertas cortas.

V. CHILE Y LOS SISTEMAS DE EXÁMENES EXTERNOS

Un punto importante para la discusión del tema es analizar lo que se hace en Chile en esta materia. El SIMCE (sistema de medición de la calidad de la enseñanza) no corresponde a un sistema de exámenes de competencias mínimas ni al de exámenes finales de cursos externos basados en el currículo. En primer lugar, no entrega resultados individuales por niño y no asocia consecuencias individuales a sus resultados. Si bien el SIMCE establece estándares generales por ciclo, ellos no obligan a alinear las calificaciones en la escuela. Por lo tanto, en enseñanza básica y media las notas de los alumnos dependen por completo de la política de los establecimientos y de los profesores. Las escuelas y liceos deciden libremente quiénes reciben el certificado de promoción y graduación en enseñanza básica y media. De esta manera el valor informativo de los certificados de graduación de enseñanza básica y media difieren en forma significativa entre una institución y otra.

La prueba de aptitud académica (PAA) actualmente vigente tampoco puede considerarse como un sistema de exámenes de competencias míni-

mas ni de exámenes finales externos. En primer lugar porque una parte importante de la prueba se orienta más a la medición de habilidades de razonamiento que a conocimientos, a excepción de las pruebas específicas rendidas por un grupo pequeño de alumnos. Estas pruebas no están al servicio de evaluar lo aprendido en enseñanza media sino que están diseñadas para seleccionar y predecir quiénes tendrán más posibilidades de cursar con éxito la educación superior. No tendrían por qué exigir cuentas de lo que los alumnos han aprendido en su paso por el colegio, a menos que se estableciera que este tipo de exámenes es un mejor predictor del desempeño de los alumnos en la universidad que la actual prueba. Por otra parte, en la PAA interesa discriminar finamente entre quienes tienen mayores habilidades y conocimientos. En cambio, los exámenes externos por lo general están referidos a criterios y su objetivo es definir claramente la meta que todos deben lograr. En tercer lugar, la posibilidad de exigir cuentas a los profesores se debilita porque la responsabilidad de la preparación en esta prueba no puede ser adscrita a un grupo reducido de profesores. Esta condición se agrava cuando además se puede argumentar que no todos los alumnos tienen habilidades académicas y que las pruebas fijan un estándar demasiado alto con el objeto de discriminar entre los más preparados.

Por otra parte, las pruebas de ingreso a la universidad actúan como un incentivo poderoso sólo para un grupo muy reducido de alumnos. De la cohorte de alumnos que ingresan a primero básico sólo un 31%⁵⁵ la rinde y un 63% de los egresados de enseñanza media se presenta a ella, considerando en el cálculo a los alumnos de enseñanza media científico-humanista y técnico-profesional. Sin embargo, alrededor de la mitad de éstos no tienen afanes competitivos reales. Aunque un número creciente de alumnos está dando la PAA cada año, muchos de ellos se presentan sin preparación porque necesitan rendirla solamente para postular a institutos o trabajos.

El nuevo sistema de ingreso a la educación superior (SIES) tampoco será un examen de competencias mínimas, ni una colección de exámenes externos finales basados en el currículo. Al igual que la prueba de aptitud académica, el SIES no está pensado para evaluar el aprendizaje de la enseñanza media sino que para seleccionar alumnos para la educación superior. Por consiguiente, no define un estándar que todos los alumnos debieran alcanzar.

⁵⁵ Dato calculado para la generación que ingresó en 1983 y rindió la prueba en 1999. Véase B. Eyzaguirre y C. Le Foulon, "La Calidad de la Educación en Cifras" (2001), p. 126.

En el marco actual de la educación en Chile, una proporción importante de los alumnos puede terminar enseñanza básica y media sin haber hecho el más mínimo esfuerzo y por el sólo hecho de haber “aparecido” en la escuela. Es necesario, entonces, hacerse cargo de la necesidad de exigir cuentas a todos los alumnos.

Conclusiones

Los alumnos no son objetos inanimados, son individuos con motivaciones y voluntad propia que pueden responder con su esfuerzo. En las culturas asiáticas está plenamente validado exigir cuentas a los alumnos porque su concepción de logro escolar implica estudio y aplicación más que habilidad. En algunos países europeos, Alemania por ejemplo, la concepción es mixta. Se considera que el talento juega un rol importante, pero al interior de cada grupo de habilidad se reconoce la importancia del esfuerzo y de acuerdo a ello exigen en forma diferenciada. En el sistema educativo norteamericano se atribuía un peso mayor a la influencia del talento en el rendimiento, lo cual se traducía en evaluaciones más personalizadas. En la última década esta visión ha sido cuestionada y se está imponiendo la idea de someter a la mayoría de los alumnos a un estándar mínimo.

La teoría indica que los sistemas de rendición de cuentas de los alumnos afectan a los incentivos de los estudiantes, de los padres y de las escuelas y, por ende, a la calidad de la educación. Las investigaciones empíricas sobre los exámenes externos con consecuencias para los alumnos, aunque no son perfectas desde un punto de vista técnico, muestran evidencia creciente de que un número significativo de alumnos aumenta su nivel de aplicación cuando se espera más de ellos.

En Europa y los países asiáticos que alcanzan un buen nivel educacional, examinar a los alumnos externamente es una práctica común. En Estados Unidos, alrededor del año 2005 todos los estados tendrán exámenes de competencias mínimas para egresar de enseñanza media, independientes de las pruebas de ingreso a la universidad, y un número creciente ha empezado a imponer exámenes externos en los cursos intermedios clave. Chile no ha implementado un sistema de rendición de cuentas individuales que afecte a todos sus alumnos, ni en las etapas intermedias de la educación escolar ni en las finales. Al menos, debiéramos colocar el tema en la agenda.

VI. PROPUESTA DE UN SISTEMA DE RENDICIÓN DE CUENTAS DE LOS ALUMNOS EN CHILE

La siguiente propuesta persigue generar discusión sobre la base de un esquema concreto. Los incentivos que generan los sistemas de rendición de cuentas son poderosos, por lo tanto, un diseño cuidadoso es esencial para que los resultados obtenidos se orienten en la dirección esperada. Lo que sigue está muy lejos de un sistema acabado, son sólo lineamientos preliminares.

6.1. Objetivos

Este sistema de rendición de cuentas de los alumnos busca elevar el estándar general al cual aspira la educación chilena. Las aspiraciones deben ir más allá de una declaración de intenciones, por consiguiente, el compromiso debe expresarse en la voluntad de exigir logros mínimos a los alumnos. El foco de la intervención se centra en elevar el rendimiento de los alumnos que no han alcanzado los objetivos mínimos y contenidos fundamentales en las asignaturas instrumentales (lenguaje y matemática). A la vez, no debiera dejar de tener expectativas altas para los alumnos que pueden alcanzar niveles mayores de aprendizaje. En este sentido tiene que reconocer los esfuerzos de los alumnos que van más allá del mínimo. La propuesta de Fontaine (2002) sería un buen complemento para lograr este último objetivo entre los alumnos de las escuelas subvencionadas.

6.2. Estándares y exámenes referidos a criterios

Los estándares deben definirse de acuerdo al currículo vigente y establecerían cuatro categorías: insuficiente, elemental, logrado y avanzado. El nivel logrado corresponde a un manejo seguro de los contenidos mínimos y objetivos fundamentales del grado. Alcanzar este nivel implicaría que el alumno podrá enfrentar los objetivos del curso siguiente sin necesidad de volver a trabajar desde sus inicios las bases anteriores. El nivel avanzado tendría que estar de acuerdo con lo que está logrando actualmente el 10% de los mejores alumnos del mundo en esa asignatura. Debe consignar excelencia académica. El nivel elemental indicaría un retraso pedagógi-

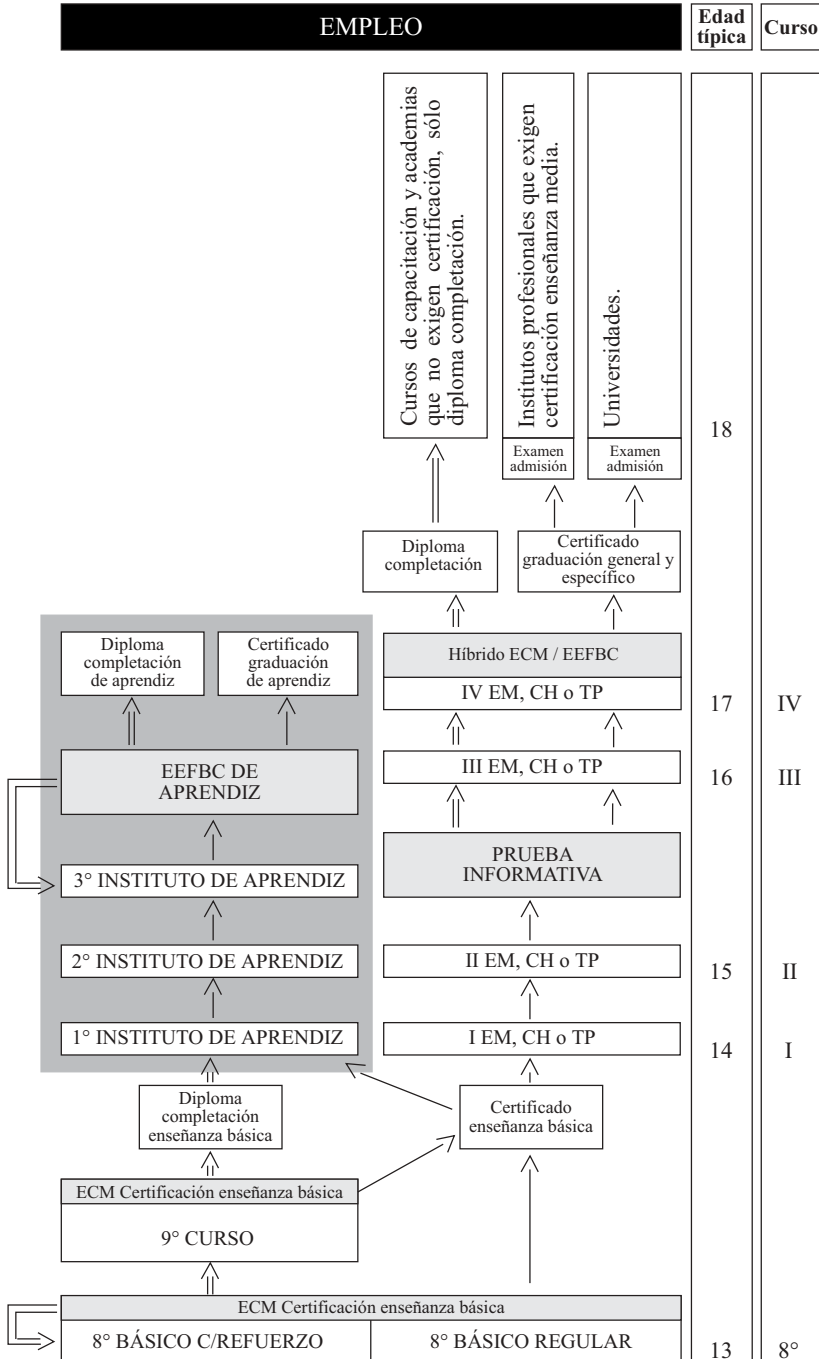
co moderado. El nivel insuficiente, que lo esperado está lejos de lograrse y que probablemente el alumno no estará en condiciones de abordar las exigencias del curso siguiente que suponen esos aprendizajes previos.

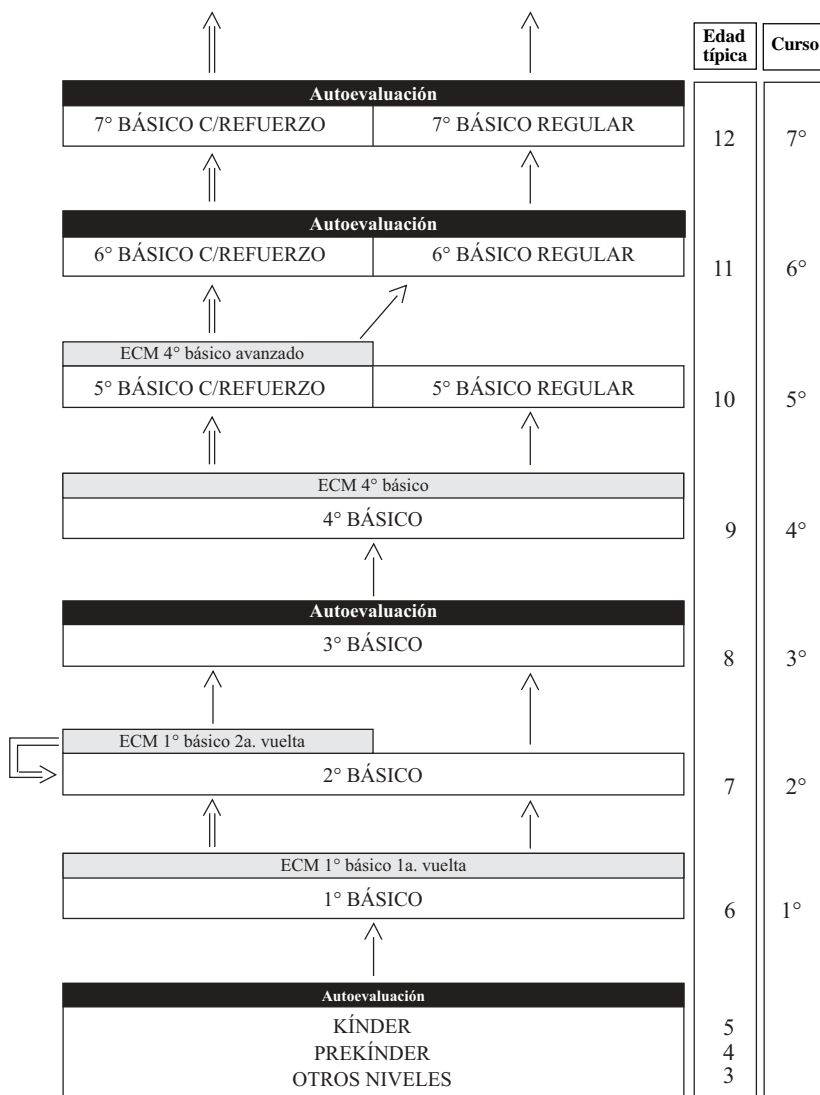
Los exámenes deberían estar referidos a criterios. Esto significa que deberían construirse en torno a estándares previamente fijados. La gran mayoría de los alumnos podría alcanzar el “nivel logrado” si se esforzaran y la calidad de la educación mejoraría. En este sentido, no es una prueba relativa que se adecua a lo que los alumnos saben para discriminar finamente entre ellos. No busca jerarquizar sino que definir un deber ser. Los objetivos de logro sobre los cuales se construyen los exámenes deberían publicarse y entregarse a todos los establecimientos, padres y alumnos en cursos superiores.

Los costos de esta propuesta se justificarían plenamente ya que se sabe que las inversiones en educación sin sistemas de rendición de cuentas, por lo general, no se acompañan de un aumento en la calidad de la educación⁵⁶. Hoxby (2002) analiza el costo de los exámenes externos aplicados en distintos estados de EE.UU. Los costos por año y por alumno fluctúan entre US\$ 2 y US\$ 25. Un sistema sofisticado como el de Texas gasta US\$ 20 por alumno al año, lo cual incluye la elaboración de las pruebas, su aplicación y corrección. Se evalúa matemática y comprensión de lectura desde 3° hasta 8° año básico y al finalizar la educación media. Además, se examina ciencias sociales y ciencias en 8° año básico y redacción en 4°, 8°, y 12° año. En enseñanza media hay pruebas para los cursos específicos de álgebra, biología, historia de Estados Unidos e inglés nivel II. Estos exámenes están totalmente vinculados a los estándares de Texas para cada curso. También se administran pruebas especiales para los estudiantes de educación diferencial y se están desarrollando exámenes para alumnos superdotados y talentosos. Su presupuesto considera la entrega de un informe individual a los padres y a las escuelas, la elaboración de boletines con información detallada para el público general y el gasto en relaciones públicas. Por último, cuenta con una base de datos completa que permite un seguimiento longitudinal de los alumnos a lo largo de su carrera escolar y universitaria.

⁵⁶ Véanse las reflexiones de Harald Beyer sobre la productividad de las inversiones en educación durante las últimas décadas en Chile, en “Algo Más sobre la Educación” (2002).

SISTEMA DE EXÁMENES EXTERNOS PARA CHILE





EM: enseñanza media.
 CH: científico-humanista.
 TP: técnico-profesional.
 ECM: examen de competencias mínimas.
 EEFBC: exámenes externos finales basados en el currículo.

↑↑ = Reprueba exámenes ↑ = Aprueba exámenes

A. Evaluación del primer ciclo de educación básica

Se propone realizar una evaluación en el primer ciclo que condicione la promoción de los alumnos a tercero básico. Esta evaluación podría estar a cargo del Ministerio de Educación o se podría estudiar la posibilidad de que haya más de una institución evaluadora como en Inglaterra. El proceso tiene que dar más de una oportunidad de rendir la prueba y para esto se sugiere el siguiente esquema: todos los alumnos rendirán un examen de lectura, escritura, y matemática inicial al final de primero básico. Los que lo reprueben tendrán todo el segundo año básico para remediar este déficit y tienen que repetir el examen al final de ese curso. Quienes no aprueban esta nueva vez deberán volver a cursar segundo. El sistema debe permitir un margen de libertad para promover a los que tienen dificultades severas y que se beneficiarán psicológica y socialmente si no se los hace repetir; por ejemplo, niños con síndrome de Down, con deficiencias cognitivas moderadas y problemas de aprendizaje, etc. Para esto, cada colegio podría usar su propio criterio para promover o retener un 10%⁵⁷ de cada generación aun cuando no haya cumplido el mínimo necesario. Esta medida facilitaría la integración y evitaría que los establecimientos discriminen a los niños que tienen problemas de aprendizaje al momento de ingresar a la escuela.

En lenguaje, el nivel elemental para segundo básico debiera ser, por ejemplo, la lectura de un mínimo de 50 palabras por minuto y la comprensión literal de textos simples, más la escritura de oraciones simples. El niño que logre este nivel en primero básico quedaría eximido de la prueba en segundo y después de cursar dicho curso podría ser promovido a tercero. Los niños que no alcancen el nivel elemental en primero deberían repetir la prueba en segundo básico. Sólo podrían ser promovidos a tercero si alcanzaron el nivel elemental. En matemática se procedería de manera similar. El nivel elemental debería exigir en segundo básico, por ejemplo, sumas con reserva y restas con canje con dos dígitos.

Se propone un número fijo de años de gracia antes que empiecen a operar las consecuencias para los niños; esto daría el tiempo necesario para que las escuelas y los padres hagan los ajustes del caso. Mientras tanto, sólo

⁵⁷ Se sugiere preliminarmente este porcentaje como un ejemplo. Para determinarlo habría que revisar la experiencia comparada, ya que la mayoría de los países que tienen exámenes externos contemplan este tipo de mecanismos. Condicionar esta franquicia a la presentación de certificados psicopedagógicos, psicológicos y psiquiátricos emitidos por especialistas es más engorroso, ya que se debería financiar el trabajo de estos profesionales y se corre el riesgo de que se produzca una inflación del número de eventos que requieran esta exención. En todo caso, en muchos países se exige incorporar en la información de los resultados de la escuela el número de alumnos eximidos.

se aplicarían las pruebas y se entregaría la información a los padres. En cuatro años un colegio podría replantear sus métodos de enseñanza y los apoderados podrían presionar a la escuela para que mejore sus resultados o bien escoger una escuela que les dé más garantías de que su hijo recibirá una buena enseñanza y apoyo.

Después de la mitad de los años de gracia, los directores, coordinadores y profesores de colegios que tengan un porcentaje de alumnos que reprobren el examen de primer ciclo básico mayor que el fijado, deberían obligatoriamente aprobar un curso práctico de enseñanza de lectura y de matemática inicial de al menos 60 horas de clase⁵⁸. El pago de esta capacitación debería ser financiado por la propia institución por intermedio del SENCE (Servicio Nacional de Capacitación y Educación) o con los fondos que se han entregado para asistencia técnica. Los organismos que impartan este curso deberían a su vez hacer públicos los resultados de sus intervenciones, de manera que los establecimientos conozcan el nivel de efectividad de su capacitación. Si la situación no se corrige después del cuarto año, habría que pensar en una medida más radical, como retirarles algunos privilegios del estatuto docente al director y coordinador académico⁵⁹, ya que de ellos depende la organización general de la escuela.

Los resultados de la escuela se comunicarían a los padres, a la escuela y a la comunidad. Los resultados individuales se entregarían a la escuela y a los padres de los respectivos niños. Se excluirían del informe de la comunidad los resultados del 10% de alumnos con peor rendimiento del curso. Como ya se dijo, esto persigue que las escuelas acepten a niños que presenten dificultad sin perjudicar los indicadores de rendimiento del establecimiento. Sin embargo, los padres de los niños que tienen dificultades recibirían el informe de resultados de sus hijos. La información al público incluiría los puntajes obtenidos en la prueba en primero básico y el porcentaje de alumnos que son promovidos a tercero básico, además de los indicadores de nivel socioeconómico y repitencias en cursos anteriores⁶⁰.

⁵⁸ Se propone esta u otra medida equivalente que genere capacidad en la escuela para enseñar lectura y cálculo inicial. Las escuelas podrían comprometerse con medidas distintas que apunten a solucionar el problema.

⁵⁹ El espíritu de esta medida es considerar los beneficios del Estatuto Docente como un privilegio que se conserva mientras los educadores cumplan adecuadamente las funciones básicas del docente. En este caso: crear las condiciones para que los niños aprendan lo elemental para iniciar su vida escolar con éxito. Esta medida fue propuesta por H. Beyer en "Falencias Institucionales en Educación: Reflexiones a Propósito de los Resultados del TIMSS" (2001).

⁶⁰ El Sistema Nacional de Evaluación del Desempeño (SNED) recoge una serie de indicadores que permiten acercar la evaluación de las escuelas a una medida de valor agregado. Incluir datos similares en la información entregada a la comunidad y los apoderados los ayudaría a tomar mejores decisiones.

B. Autoevaluación de las escuelas en 3^{er} año básico

El Ministerio de Educación debiera elaborar o adaptar pruebas estandarizadas referidas a criterio en las diferentes asignaturas para que el establecimiento autoevaluara el cumplimiento de los objetivos de cada curso. Cada escuela tendría que aplicarlas, corregirlas y tabularlas⁶¹. La información que ellas entregarían sería de uso interno y permitiría a la dirección de la escuela monitorear el aprendizaje. También ayudaría a estimar la probabilidad de alcanzar las metas establecidas en los exámenes posteriores. Este proceso contribuiría a que la dirección y los profesores internalizaran los estándares esperados y a organizar el trabajo en torno a objetivos concretos. Eventualmente, si la escuela lo estimara conveniente, podrían informar los resultados a los padres y alumnos.

Esta medida también se sugiere para al término de kínder y todos los cursos intermedios que no son examinados externamente.

C. Examen final de segundo ciclo (al término de 4° año básico)

Los alumnos rendirían un examen al final de 4° básico, en lo fundamental siguiendo las condiciones del que rindieron en el primer ciclo básico. Las variantes que se introducen son las siguientes: los alumnos que no alcancen el nivel elemental estipulado obligatoriamente deberían remediar su retraso pedagógico en un período determinado de apoyo complementario. La calendarización y el horario se pactarían con el establecimiento. Entre las alternativas están el verano, después de la jornada escolar o los días sábados⁶². De ninguna manera en el horario regular. El colegio debería asignar parte de la subvención para cubrir los costos de esta medida remedial. El Mineduc velaría para que se cumpla esta disposición. Al final de quinto básico estos alumnos rendirían nuevamente la prueba de 4° básico y tendrían que alcanzar el “nivel logrado” para no seguir con apoyo remedial en sexto básico. Si no lo aprobaran, los alumnos deberían continuar con apoyo en sexto y séptimo.

D. Autoevaluación para 5°, 6° y 7° básico

Esta instancia se refiere a la descrita en la letra B.

⁶¹ Habría que estudiar si sería conveniente repartirlas centralizadamente o si cada colegio tendría que reproducirlas.

⁶² Véase la experiencia de Chicago descrita en R. P. Phelps, “Why Testing Experts Hate Testing” (1999).

E. Examen final de enseñanza básica

A mediados de octubre, los alumnos rendirían un examen de competencias mínimas en lectura, redacción y matemática⁶³. Para aprobarlo se exigiría pasar el nivel elemental, lo cual daría derecho a la certificación de la enseñanza básica y al ingreso a educación media. Quienes no lo aprueben podrían repetirlo en diciembre y otra vez en enero. Si después de estos ensayos el alumno no lo lograra, tendría que repetir curso en su escuela o asistir a un 9° grado que se concentraría en remediar déficits en lenguaje y matemática. Este curso lo ofrecerían las municipalidades o escuelas que se interesaran en ello y recibirían subvención al igual que los cursos regulares. Al término de 9° grado los estudiantes rendirían nuevamente el examen. Aquellos que no logran aprobarlo recibirían un diploma de completación de la enseñanza básica⁶⁴. Los alumnos que estimaran conveniente quedarse con este diploma sin rendir 9° año podrían hacerlo.

Para los jóvenes que logren el diploma de completación se sugiere estudiar la posibilidad de crear institutos de aprendices subvencionados, al igual que la enseñanza media, con una duración de dos a tres años. La subvención que habrían recibido los alumnos al seguir sus estudios en la enseñanza regular se trasladaría a estos nuevos institutos que serían regulados por el Ministerio de Educación. En ellos los alumnos lograrían una capacitación de buen nivel pero más práctica. Esto apuntaría a resolver el problema de los alumnos que no tienen inclinaciones o aptitudes para el estudio académico o el técnico-profesional más avanzado. Seguramente parte de estos jóvenes son los que actualmente desertan entre octavo básico y segundo medio.

Los resultados obtenidos en este examen final de enseñanza básica se comunicarían con la misma prontitud que los de la prueba de aptitud académica, para que los estudiantes que no lo pasaron en la primera instancia se preparen para darlo nuevamente. Para lograr la corrección en el plazo estipulado, la prueba de redacción se corregiría internamente en las escuelas con pautas preestablecidas. Para asegurar la confiabilidad de este procedimiento, los alumnos redactarían sus composiciones con copia⁶⁵, una de las cuales quedaría en manos del colegio y la segunda sería retirada junto

⁶³ Ciencias sociales y ciencias naturales podrían seguir aplicándose pero no se considerarían en un principio para retener el certificado de enseñanza básica. En el futuro estas asignaturas podrían integrarse como un requisito.

⁶⁴ La posibilidad de rendir el examen quedaría abierta para siempre previo pago de su costo. Los adultos que completen su educación también tendrían la posibilidad de rendirlo.

⁶⁵ El examinador externo entregaría el tema de la redacción en hojas autocalcantes. La escuela se quedaría con la copia y el ministerio con el original.

con el resto de las pruebas. Se haría un chequeo al azar y se permitiría una desviación determinada; en caso de comprobar una transgresión se daría paso a una amonestación. La posibilidad de corregir las propias pruebas es una instancia de desarrollo pedagógico importante para los profesores.

Los puntajes generales se publicarían después de la última evaluación y se consignaría la proporción de alumnos que aprobaron junto con el número de alumnos que se posicionaron en cada uno de los niveles (insuficiente, elemental, logrado y avanzado). Al igual que en las instancias anteriores, la comunicación de los resultados a la comunidad y a los padres será fundamental.

F. Autoevaluación de I medio

Esta instancia se refiere a la descrita en la letra B.

G. Prueba Informativa de II medio

La prueba informaría a los alumnos, padres y escuela el nivel de logro alcanzado en los dos primeros años de educación media. Con los resultados se determinaría un índice que permitiría calcular las posibilidades de aprobar los exámenes mínimos al final de IV medio si el alumno sigue con el mismo ritmo de estudio. La medida señalaría al alumno la necesidad de cambiar sus hábitos de trabajo⁶⁶. Los resultados se publicarían igual que en los casos anteriores. No se imponen consecuencias adicionales.

H. Exámenes finales de enseñanza media

El sistema de examinación de IV medio es un híbrido compuesto por exámenes de competencias mínimas y una colección de exámenes externos finales basados en el currículo. Los alumnos recibirían un “certificado general de enseñanza media” si aprueban lo exigido en el examen de competencias mínimas; en caso de que no lo aprobaran se entregaría un diploma de completación de enseñanza media. El sistema reconocería, además, a todos aquellos alumnos que logren alcanzar conocimientos y destrezas que van más allá de los mínimos exigidos para graduarse de la enseñanza me-

⁶⁶ El sistema diseñado por el Departamento de Educación del estado de Texas contempla un indicador de este tipo. Véase, Texas Education Agency, “Accountability Manual: The Accountability Rating System for Texas Public Schools and School Districts” (2000).

dia. Con este fin los estudiantes podrían optar por un “certificado de enseñanza media específico”.

El certificado general de enseñanza media es un examen que deberían rendirlo todos los alumnos y demandará para IV año de enseñanza media un buen dominio de las destrezas generales exigidas para II medio (comprensión de lectura y razonamiento matemático) y en contenidos cubrirá lo más elemental hasta IV año de enseñanza media. Para definir bien los estándares a exigir en esta prueba, sería necesario revisar el actual currículo nacional de enseñanza media y acotar lo que se considere elemental. El currículo actual define contenidos mínimos y objetivos fundamentales, pero en la práctica, en algunas asignaturas, éstos exceden lo que se puede considerar mínimo. Básicamente esto sucede porque no están redactados como objetivos de logro, lo cual se traduce en una definición demasiado abierta de lo que se quiere lograr. Se rendirían en cinco asignaturas: matemática, lenguaje, ciencias naturales y sociales, e inglés⁶⁷. Todas las asignaturas considerarían destrezas y conocimientos, a excepción de lenguaje, que abordaría fundamentalmente la medición de las destrezas de comprensión lectora y redacción. Serían pruebas referidas a criterio que distinguirían tres niveles: insuficiente, elemental y logrado. El Estado financiaría hasta tres oportunidades de rendición; otras posibilidades tendrían que ser costeadas por los propios alumnos y estarían abiertas para todos aquellos que deseen certificar su enseñanza media. Los exámenes se darían una vez al año y los alumnos que quisieran podrían rendirlos a partir de III medio.

Los exámenes de certificación específicos son colecciones de exámenes referidos a criterios en temas específicos de cada asignatura. Señalan niveles de logro más avanzados que los que fijan los exámenes de competencias mínimas. A grandes rasgos, el modelo seguiría el esquema de Nueva Zelanda⁶⁸. Los alumnos tendrían que completar un mínimo de créditos en cada una de las dos grandes áreas evaluadas, humanidades y ciencias, y un número de créditos adicionales en una de ellas.

Los créditos se acumularían aprobando los requisitos exigidos en cada uno de los exámenes específicos que los alumnos decidan rendir. Estos abarcarían temas de cada una de las asignaturas. Los contenidos se desprenderían del plan general de formación y del plan diferenciado del currículo nacional. Por ejemplo, en biología podríamos encontrar una pue-

⁶⁷ Inglés exigiría un nivel muy básico para aprobar. El examen se haría cargo, por una parte, de que todos los alumnos han tenido ocho años de inglés obligatorio y, por otra, que el país aún no tiene profesores de alto nivel para todas sus escuelas.

⁶⁸ Véase National Certificate of Educational Achievement of New Zealand, “National Qualifications Framework” (www.nzqa.govt.nz/framework/about/index.html).

ba sobre célula, genoma y organismo y otra sobre evolución, ecología y ambiente. Lo mismo en historia, por ejemplo, una podría estar centrada en “La ciudad contemporánea” o “El siglo XX y la búsqueda del desarrollo económico y la justicia social”. Los estudiantes de educación técnico-profesional tendrían que escoger exámenes específicos en su área.

El número de créditos dependería del nivel de dificultad de cada examen. Además cada prueba tendría un puntaje que señalaría distintos grados de dominio. Así, los alumnos que rindieran estas pruebas tendrían un certificado que señalaría que se ha aprobado el mínimo de créditos exigidos y una transcripción con los créditos aprobados y los puntajes obtenidos en cada evaluación.

Los exámenes específicos podrían reemplazarse por el Bachillerato Internacional u otro de igual prestigio, si son de un nivel de dificultad equivalentes o más elevados que los exámenes específicos nacionales. La equivalencia la determinaría el organismo que diseña el resto de las pruebas o un consejo consultivo.

Los específicos se podrían rendir una vez al año y sólo la primera vuelta sería gratuita. Los alumnos que quisieran rendirlos anticipadamente podrían hacerlo desde III año de enseñanza media.

Certificado general de enseñanza media:

Certificado específico de enseñanza media

Matemática: plan de formación general.

Matemática: plan general y diferenciado (matemática avanzada, geometría para diseño, estadística para ciencias sociales, etc.).

Lenguaje: plan de formación general (comprensión lectora y redacción)

Literatura: elección de obras de un período de la literatura universal o de la literatura latinoamericana, etc.

Historia y ciencias sociales: plan de formación general.

Historia y ciencias sociales: elección entre posibilidades de temas específicos; por ejemplo, historia del arte.

Ciencias: plan de formación general.

Biología o física o química: nivel diferenciado.

First Certificate de Cambridge, o Toefle u otro equivalente.

Idioma extranjero (inglés): plan de formación general.

Los alumnos de técnico-profesional tendrían que elegir temas de su sector: construcción, eléctrico, maderero, agropecuario, etc.

La aprobación de todos los exámenes rendidos quedaría escrita en una transcripción que recibiría todo alumno al graduarse. Ésta incluiría el promedio de notas general de la enseñanza media y los créditos y puntajes en cada una de los exámenes externos finales rendidos. Esta transcripción se podría utilizar como información para el ingreso a una universidad, si alguna de ellas lo estima conveniente, o para el ingreso a puestos de trabajo. Los puntajes y proporción de alumnos que reciben el certificado de educación media se publicarían igual que en los casos anteriores.

El sistema de exámenes esbozado en este trabajo busca generar poderosos incentivos al estudio durante todo el período escolar, desde el primer ciclo básico hasta el término de la educación media. Cabe señalar que en la discusión sobre la reformulación del sistema de admisión a la universidad se ha planteado que el proyecto “Sistema de Ingreso a la Educación Superior” (SIES) también cumpliría con el propósito de crear los incentivos al estudio que hacen falta en la educación media. Sin embargo, no se pueden esperar los mismos efectos de uno y otro sistema. En primer lugar, los ECM exigen cuentas a todos los alumnos de enseñanza media, independientemente del interés que ellos manifiesten por continuar estudios superiores, en cambio en Chile las pruebas de ingreso a las universidades son optativas. En 1999, sólo un tercio de la cohorte de los alumnos que ingresó a 1° básico y aproximadamente dos tercios de los egresados de la educación media rindieron la prueba de admisión a la universidad. Es más, un porcentaje importante de los que la dan sólo lo hacen para acreditar prueba rendida para presentarse al trabajo, y entre los que aspiran a la universidad hay quienes no harán esfuerzos extraordinarios porque se conforman con carreras no selectivas. Para este grupo, y para el 40% que no rinde la prueba, los incentivos del SIES no operarán o lo harán débilmente.

Por otro lado, en lo que concierne al objetivo de mejorar la calidad de la educación escolar, un sistema que exige cuentas desde los primeros años de educación básica constituye un vigoroso instrumento, porque envía señales tempranas que permiten corregir el rumbo antes que se acumulen los déficits. En cambio, las pruebas de ingreso a la universidad operan tarde, cuando los padres y los alumnos poco pueden hacer para remediar su rendimiento.

Otra diferencia importante con las nuevas pruebas de ingreso a la universidad consiste en la libertad pedagógica que permite el esquema sugerido en este trabajo. Para comprender por qué los exámenes de competencias mínimas y los exámenes externos finales basados en el currículo dejan mayores márgenes de movilidad es necesario entender las diferencias que existen entre las evaluaciones destinadas a la selección de alumnos y

aquellas que fijan un estándar que puede ser aprobado por todo aquel que logre lo esperado⁶⁹. En pruebas selectivas como las de ingreso a la universidad lo más importante es jerarquizar a los alumnos finamente en el parámetro que se considera predictivo del éxito universitario. Por lo tanto, se construyen de manera tal que sólo unos pocos pueden obtener los máximos puntajes. En ellas, no son las preguntas promedio las que determinan un puntaje alto, sino que contestar lo que nadie sabe. Por ello es que las pruebas de ingreso a la universidad centradas en contenidos terminan transformándose, en la práctica, en un currículo máximo que no deja espacio a los programas propios porque los establecimientos deben obtener una ventaja sobre los demás, la que se materializará estudiando cada minucia si el currículo es restringido o tratando de lograr la máxima cobertura si el currículo es amplio. En cambio, las pruebas referidas a estándares como los ECM fijan claramente las metas esperadas, y una vez que los establecimientos las han logrado quedan en libertad para perseguir los objetivos complementarios que han fijado en sus propios proyectos educativos. Así, el esquema presentado provee tempranamente incentivos para todos y respeta la necesaria libertad de educación.

6.3. Consideraciones finales

Los incentivos que establece un sistema de rendición de cuentas como el propuesto son poderosos. Un cambio de esta naturaleza requiere una discusión informada y mucha investigación para llegar a un buen diseño. La experiencia comparada es vasta y de ella se podrían extraer lecciones valiosas. Esta proposición es sólo un punto de partida.

REFERENCIAS BIBLIOGRÁFICAS

- Aguilar, M. "Tiempos Violentos: Se Cierra un Agitado Año Escolar." *El Mercurio*, 14 de diciembre, 2001.
- American Educational Research Association (AERA). "AERA Position Statement Concerning High-Stakes Testing in PreK-12 Education". Aera.net, 2001.
- Betts J. R., y Costrell, R. M. "Incentives and Equity under Standards-Based Reform". En D. Ravitch (ed.), *Brookings Papers on Education Policy 2001*. Brookings Institution Press, 2001.
- Beyer, H. "Educación y Desigualdad de Ingresos: Una Nueva Mirada". *Estudios Públicos* N° 77, 2000.

⁶⁹ Para una discusión más amplia de este punto, véase B. Eyzaguirre y C. Le Foulon, "El SIES: Un Proyecto Prematuro" (2002).

- Beyer, H. "Falencias Institucionales en Educación: Reflexiones a Propósito de los Resultados del TIMSS". *Estudios Públicos*, N° 82, 2001.
- Beyer, H., Eyzaguirre, B., y Fontaine, L. "La Reforma Educacional Chilena, editado por Juan Eduardo García Huidobro". (Reseña de libro.) *Perspectivas en Política, Economía y Gestión*, Vol. 4, N° 2, Universidad de Chile, 2001.
- Beyer, H. "Falencias Institucionales en Educación: Reflexiones a Propósito de las Reflexiones del TIMSS". *Estudios Públicos*, 82, 2001.
- Beyer, H. "Algo Más sobre Educación". Mimeo, Centro de Estudios Públicos, 2002.
- Bishop, John H. "Impact of Academic Competency on Wages, Unemployment, and Job Performance". *Carnegie-Rochester Conference Series on Public Policy*, Vol. 37, diciembre 1992.
- Bishop, J. H. "The Impact of Curriculum-Based External Examinations on School Priorities and Student Learning". *International Journal of Educational Research*, Vol. 23, N° 8 (1996).
- Bishop, J. H. "The Effect of National Standards and Curriculum-Based Exams on Achievement". *The American Economic Review Papers and Procedures*, Vol. 87, N° 2, 1997.
- Bishop, J. H. "The Effect of Curriculum-Based External Exit Exam Systems on Student Achievement". *Journal of Economic Education*, Vol. 29, 2 primavera 1998.
- Bishop, J. H., y Mane, F. "The Impact of Minimum Competency Exam Graduation Requirements on College Attendance and Early Labor Market Success of Minority Students". Civil Right Project, Harvard University, diciembre, 1998. www.law.harvard.edu/civilrights/conferences
- Bishop, J. H., Mane, F., y Bishop, M. "La Educación Secundaria en Estados Unidos: ¿Qué pueden aprender otros de nuestros errores?". Center for Advance Human Resource Studies, Cornell University, 2000a.
- Bishop, J. H., Moriarty, J., y Mane, F. "Diplomas for Learning, Not for Seat Time: The Impact of New York Regents Examinations". *Economics of Education Review*, 19, 2000b.
- Bishop, J. H., y Wößmann. "Institutional Effects in a Simple Model of Educational Production". Kiel Institute of World Economics, Alemania, noviembre, 2001.
- Bishop, J. H., Mane, F., Bishop, M., y Moriarty, J. "The Role of End of Course Exams and Minimum Competency Exams in Standards-Based Reforms". En D. Ravitch (ed.), *Brookings Papers on Education Policy 2001*. Brookings Institution Press, 2001.
- Boser, U. "Pressure Without Support". Education Week on the Web, 2001.
- Britton, E. D., y Raizen. S. A. *Examining the Examinations. An International Comparison of Science and Mathematics Examinations for College Bound Students*. Londres: Kluwer Academic Publishers, 1996.
- Catterall, J. S., y Winters, L. "Economic Analysis of Testing: Competency, Certification, and 'Authentic' Assessments". CSE Technical Report 383, CRESTT, UCLA, 1994.
- Center for Education Research and Innovation. *Schools under Scrutiny*. París: OECD, 1995.
- CIDE. "II Encuesta Nacional a los Actores del Sistema Educativo 2000". Santiago: CIDE, 2001.
- Cizek, G. J. "Unintended Consequences of High Stakes Testing". Draft (Informe Preliminar) School of Education, University of North Carolina, 2001.
- Clarke, M., Haney, W., y Madaus, G. "High Stakes Testing and High School Completion". NBETPP (National Board on Educational Testing and Public Policy), Boston College, Vol. 1, N° 3, 2000.

- Costrell, R. M. "A Simple Model of Educational Standards". *The American Economic Review*. Vol. 84, N° 4, septiembre 1994.
- Costrell, R. M., y Ardon, K. "MCAS and the Rise of Literacy Skills in the Early Grades, 1998-1999". The Commonwealth Office, Report Series N° 6, 2000.
- Deere, D., y Strayer, W. "Putting Schools to the Test: School Accountability, Incentives, and Behavior". Texas A&M University, 2001.
- Donoso, G., Bocchieri, M. A., Ávila, E., Núñez, N. *et al.* *El Sistema de Admisión: Orígenes y Evolución*. Santiago: Departamento de Evaluación, Medición y Registro Educativo (DEMRE), Universidad de Chile, 2000.
- Education Week*. "Making the Grade: Teachers' Attitudes Toward Academic Standards and State testing. Findings of National Survey of Public School Teachers". Education Week on the Web, noviembre 2000.
- Education Week*. "Standards Related Tables. Quality Counts 2001: A Better Balance". Education Week on the Web, 2001.
- Elliot, J. L., y Hyeonsook, S. A. "Perspective on Education and Assessment in Other Nations: Where are Students with Disabilities?". NCEO (National Center on Educational Outcomes), 1995 (<http://education.umn.edu/nceo>).
- Eysenk, H. J. *A New Look to Intelligence*. New Jersey: Transaction Publishers, 1998.
- Eyzaguirre, B., y Le Foulon, C. "La Calidad de la Educación en Cifras". *Estudios Públicos*, N° 84, 2001.
- Eyzaguirre, B. "El SIES: Un Proyecto Prematuro". *Estudios Públicos*, N° 87, 2002.
- Farber, H., y Gibbons, R. "Learning and Wage Dynamics". *Quarterly Journal of Economics*, 1996.
- Feuerstein, R. *Instrumental Enrichment: An Intervention Program for Cognitive Modifiability*. Scott, Foresman and Company, 1980.
- Figlio N. D. "Do High Grading Standards Affect Student Performance?". NBER Working Paper N° 7985, 2000.
- Fontaine, A. "Equidad y Educación: Cinco Proposiciones". *Puntos de Referencia* N° 255, 2002, Santiago, Centro de Estudios Públicos.
- Fontaine, L., e Eyzaguirre, B. "Una Estructura que Presione a las Escuelas a Hacerlo Bien". En H. Beyer y R. Vergara (eds.), *¿Qué Hacer Ahora? Propuestas para el Desarrollo*. Santiago: Centro de Estudios Públicos, 2001.
- Fredericksen, N. "The Influence of Minimum Competency Tests on Teaching and Learning". Princeton, Educational Testing Service, 1994.
- Glennester, H. "United Kingdom Education (1997-2001)". Case paper 50, November 2001, Center for Analysis of Social Exclusion, London School of Economics.
- Grubb, N. "The Varied Economic Returns to Postsecondary Education". *Journal of Human Resources*, 28 (2).
- Harris, J. R. *The Nurture Assumption: Why Children Turn the Way They Do*. Touchstone, 1999.
- Hoxby, Caroline. "Peer Effects in the Classroom: Learning from Gender and Race Variation". NBER Working Paper Series 7867, agosto 2000.
- Hoxby, C. "The Cost of Accountability". NBER Working Paper Series, 2002.
- Jacob, B. A. "Getting Tough? The Impact of High School Graduation Exams". Harris Graduate School of Public Policy Studies, University of Chicago, 2000.
- Johnson, K. A. "The Peer Effect On Academic Achievement Among Public Elementary School Students". Heritage Foundation, mayo 2000.

- Kindermann, T. A. "Natural Peer Groups as Contexts for Individual Development: The Case of Children Motivation in School". *Developmental Psychology*, 29, 1993.
- Kulik, James A. y Kulik, Chen-Lin. "Effects of Accelerated Instruction on Students". *Review of Educational Research*, Vol. 54, N° 3, 1984.
- Ladd, Helen F. (ed.). *Holding Schools Accountable: Performance-Based Reform in Education*. The Brookings Institution, 1996.
- Landgraff, K. M. "Using Assessments and Accountability to Raise Student Achievement". ETS.org., 2001.
- Lerner, B. "Good News about American Education". *Commentary*, Vol. 91, N° 3, 1990.
- Lillard, Dean; y De Cicca, Philip. "Higher Standards, More Dropouts: Evidence Within and Across Time". Por aparecer en *Economics and Education Review*.
- Marion, S. F., Sheinker, A. "Issues and Consequences for State Level Minimum Competency Testing Programs". National Center of Educational Research, 1999.
- Marzano, R. J. *Models of Standards Implementation: Implications for the Classroom*. Colorado: McREL, 1998.
- Ministry of Education of Singapore. Cambridge Syndicate Exams. www.moe.edu.sg/exams/2002
- Monk. "Subject Area Preparation of Secondary Mathematic and Science Teachers and Student Achievement". *Economics and Education Review*, Vol. 13, N° 2, 1994.
- NGA Center for Best Practices. Graduation Exit Exam Matrix. NGA, 2001.
- Phelps, R. P. "Why Testing Experts Hate Testing". Thomas Forham Foundation, 2002. www.edexcellence.net/library/phelps.htm
- Phelps, R. P. "Test Bashing". Parte 3ra The Education Press's Cop-Out on Student Testing. Education News Org., 2002.
- Ravitch, D. (ed.). "Beep Conference on National Standards". The Brookings Institution Brown Center on Education Policy, mayo 15-16, 2000.
- Richman, S. Parent. "Power: Why National Standards Won't Improve Education". *Policy Analysis* N° 396, abril 26, 2001.
- Roderick, M. "Ending Social Promotion: Results from the First Two Years". Chicago, ILL.: Consortium on Chicago School Research, 1999.
- Betts, J. R., Costrell, R. M. "Incentives and Equity under Standard Based Reform". *Brookings Papers on Education*. Brookings Institution, 2001.
- Sherif, M., Harvey, O. J., White, B. J., y Hood, W. R. *Intergroup Cooperation and Competition: The Robbers Cave Experiment*. Norman, OK: University Book Exchange, 1961.
- Steinberg, L., Brown, B., y Dormbusch, S. M. *Beyond the Classroom*. Simon and Schuster, 1996.
- Stevenson, H., y Stigler, J. *The Learning Gap: Why Our Schools Are Failing and What We Can Learn from Japanese and Chinese Education*. Nueva York: Summit Books, 1992. [Traducción al castellano de los capítulos 7, 8 y 9, en *Estudios Públicos*, 76, 1999.]
- Tucker, M.S. y Coddling, J. B. *Standards for Our Schools. How to set Them, Measure Them and Reach Them*. San Francisco: Jossey-Bass Publishers, 1998.
- Unicef. "Inclusión de Niños con Discapacidad en la Escuela Regular". Unicef, diciembre, 2001.
- Weinstein, N. D. "Unrealistic Optimism about Future Life Events". *Journal of Personality and Social Psychology*, Vol. 39 N° 5, 1980.

- Winfield, L. F. "School Competency Testing Reforms and Student Achievement: Exploring a National Perspective". *Educational Evaluation and Policy Analysis*, 12, 1990.
- Wößmann, L. "Schooling, Resources, Educational Institutions, and Student Performance: The International Evidence". Kiel Working Paper 983, Kiel Institute of World Economics, 2001.