



Educación

ISSN: 0379-7082

reveedu@gmail.com

Universidad de Costa Rica

Costa Rica

Contreras Montes de Oca, Ileana

El papel de las pruebas nacionales en la problemática de la educación matemática: análisis de las perspectivas de los grupos responsables de su conducción

Educación, vol. 27, núm. 1, 2003, pp. 115-133

Universidad de Costa Rica

San Pedro, Montes de Oca, Costa Rica

Disponible en: <http://www.redalyc.org/articulo.oa?id=44027110>

- ▶ Cómo citar el artículo
- ▶ Número completo
- ▶ Más información del artículo
- ▶ Página de la revista en redalyc.org

redalyc.org

Sistema de Información Científica

Red de Revistas Científicas de América Latina, el Caribe, España y Portugal
Proyecto académico sin fines de lucro, desarrollado bajo la iniciativa de acceso abierto

EL PAPEL DE LAS PRUEBAS NACIONALES EN LA PROBLEMÁTICA DE LA EDUCACIÓN MATEMÁTICA: ANÁLISIS DE LAS PERSPECTIVAS DE LOS GRUPOS RESPONSABLES DE SU CONDUCCIÓN

Ileana Contreras Montes de Oca

Recibido 10-XII-2002 • Aceptado 11-III-2003

1. Introducción

El propósito de este artículo es com-partir el análisis realizado a aquellos re-sultados que tienen relación con el papel de las pruebas nacionales de la educación matemática del país, obtenidos como pro-ducto de un trabajo de investigación en el que se identificaron las perspectivas de los grupos responsables de la conducción de la educación matemática del país en cuanto a su problemática y posibles determinantes.

Se inicia con una descripción sucinta de los aspectos principales del estudio en referencia, rescatando aquellos elementos importantes para comprender el sentido de los resultados que se desean comentar.

Luego se expone el referente concep-tual en el que se enmarca el análisis. Este contiene la discusión de algunas distincio-nes básicas de los términos empleados. Una revisión de los esfuerzos y experiencias re-lativas al tema en discusión, realizados en el contexto internacional, y de los están-dares propuestos para la evaluación de la edu-cación matemática en el nivel escolar.

Seguidamente, se discuten los resul-tados del estudio mencionado, los cuales dan cuenta de las perspectivas de los for-madores de educadores en educación ma-temática, los asesores ministeriales, nacio-nales y regionales de matemáticas, y de dos grupos de profesores de matemática en

Resumen: En el artículo se analizan las perspecti-vas de los grupos responsables de la conducción de la educación de las matemáticas del país, en el nivel se-cundario, respecto al papel que han tenido las prue-bas nacionales en su problemática actual. La identifi-cación de tales perspectivas se obtuvo como producto de un trabajo de investigación cuyo objetivo inicial no incluyó a las pruebas nacionales como categoría de es-tudio, pero en el que el papel de éstas aparece como un factor relacionado con la problemática analizada. Se describe brevemente el trabajo de investigación que generó los resultados objeto de examen. El análisis se realiza utilizando como referente los estudios y posi-ciones emitidas en otros contextos, respecto a la utili-zación de pruebas nacionales y al empleo de modelos de evaluación similares al nuestro. Se señalan algu-nas fortalezas y debilidades del uso de las pruebas na-cionales y se anotan algunas recomendaciones que se derivan del trabajo.

Palabras clave: Evaluación, Pruebas Estandarizadas, Documentación de los Aprendizajes, Educación, Matemáticas, Educación Secundaria.

servicio, en relación con el papel de las pruebas nacionales en la problemática de la educación matemática del país.

Finalmente, se expone el análisis, de donde se derivan algunas conclusiones, en términos de los retos, posibilidades, y orientaciones principales que deben atenderse, en los esfuerzos que se desarrollen para valorar la calidad de la educación nacional. Tales orientaciones responden al interés de que tales esfuerzos se realicen de manera que contribuyan a promover el aprendizaje, y que eviten aquellas consecuencias perjudiciales no previstas, que han sido documentadas en otros contextos, las cuales concuerdan en gran medida, con las perspectivas de los representantes de los grupos nacionales responsables de la educación matemática en el país, consultados.

2. Características del estudio relevantes para contextualizar los resultados a compartir

El trabajo de investigación de donde se extraen los resultados relativos al contexto costarricense, los cuales serán objeto de análisis en esta presentación, se desarrolló en el período comprendido entre los años 1998 y 2001, (Contreras, 2002) y entre sus objetivos tenía los siguientes:

2.1 Contrastar las perspectivas de los representantes de los grupos responsables de la conducción de la educación matemática en el país, respecto a:

- Los principales problemas que aquejan a la educación matemática y los principales factores que los explican.
- Sus propuestas respecto a las alternativas que deben considerarse para la superación de los problemas identificados.

- La necesidad de identificar criterios que cuenten con un alto nivel de consenso, que permitan “operacionalizar” las expectativas existentes respecto al trabajo docente de los profesores de matemática de la educación secundaria.

2.2 Contrastar los criterios expresados –explícita o implícitamente– para precisar las expectativas de los participantes en relación con el trabajo docente, con aquellos que se derivan de las políticas curriculares y del quehacer educativo nacional, y con los planteamientos y teorías que al respecto se vienen formulando en el contexto internacional.

Las técnicas y procedimientos utilizados para alcanzar dichos objetivos fueron las siguientes: investigación bibliográfica y documental; consulta a representantes de los tres grupos que se identificaron como los directamente responsables de la educación matemática en el país:

- Formadores de docentes para la educación matemática e investigadores en educación matemática.
- Asesores ministeriales, nacionales y regionales, de matemática.
- Profesores en ejercicio.

La consulta se hizo mediante talleres convocados para: analizar la problemática de la educación matemática del país; definir las condiciones requeridas para superarla; y discutir la conveniencia o no de identificar los criterios para orientar el trabajo docente en este campo, en el nivel secundario. Durante el desarrollo de los talleres se realizaron entrevistas individuales y del tipo grupo focal, que fueron audio-grabadas y posteriormente sometidas a un

análisis de contenido utilizando el procedimiento estándar.

En los talleres con formadores de docentes e investigadores, participaron 33 integrantes de las unidades académicas de las universidades públicas y privadas que ofrecen la carrera de enseñanza de la matemática o que realizan investigación en este campo. Se trabajó con los que acudieron a la invitación.

Para los talleres con los asesores ministeriales el señor Vice Ministro de Educación Pública en ejercicio, envió invitación a todos los asesores nacionales y regionales de matemáticas. Se trabajó con los 21 asesores que asistieron.

En los talleres con los profesores de matemáticas en ejercicio, se trabajó con dos grupos de diez profesores cada uno, que se encontraban laborando en diversas instituciones públicas del área metropolitana.

3. Referente Conceptual

3.1 Conceptos y términos involucrados : algunas distinciones básicas

En español existen dos términos que en el vocabulario de algunas autoridades educativas, educadores y estudiantes, así como en el público general, se utilizan como equivalentes: evaluación y medición. Ambos se relacionan con la utilización de pruebas –o “tests”–, ya que con frecuencia hacen uso de ellas.

Sin embargo, tales términos no son sinónimos. Más aún, en inglés existe otro término, “assessment”, que tampoco es sinónimo de los anteriores, pero que en el idioma español no cuenta con un término específico, por lo que se suele utilizar el de “evaluación” para referirse a los significados de ambos términos usados en el inglés.

Es necesario, por lo tanto, iniciar esta exposición con una discusión que contribuya a clarificar el significado de esos términos y

la forma en que se emplean en los países con mayor tradición en este campo. En este trabajo se usa la expresión “documentación para analizar los aprendizajes”, para referirse a lo que en inglés se designa con el término “assessment” para facilitar la diferenciación.

Para referirnos a medición, utilizaremos la definición de diccionario, que en criterio de Keeves (1997), es apropiada para la mayoría de sus aplicaciones en medición educativa, y que se entiende como la asignación de una cantidad numérica para expresar la comparación de una característica respecto a una escala.

En consecuencia, las pruebas de rendimiento se utilizan para medir la cantidad de conocimiento aprendido u olvidado. Los ítemes que se utilizan en las pruebas son manifestaciones de una variable latente, no la característica en sí misma. La teoría de respuesta del ítem reconoce esta distinción, y busca estimar el aprovechamiento en la característica latente que subyace en los ítemes.

Asimismo, la teoría clásica de pruebas, también establece la diferencia entre un “valor verdadero” y un “valor crudo”. (Keeves, 1997). Pero, lo que interesa subrayar aquí, es que la medición raramente se considera como un fin en sí misma. Es simplemente una operación útil en el proceso de evaluación.

En general, y también acuerdo con Keeves (1997), el uso del término “evaluación” se reserva para aplicarlo a entidades abstractas, como programas, currícula, y situaciones organizativas. Su empleo se hace para referirse a un juicio de valor.

La evaluación generalmente requiere la realización de comparaciones contra un estándar; o contra criterios derivados de los objetivos establecidos, o de otros programas, currículum o situaciones organizacionales. La evaluación es una actividad fundamental de la investigación y el desarrollo. Puede requerir la medición de resultados educativos, y de la aplicación de

pruebas a individuos o grupos. Su potencial importancia para el mejoramiento de la práctica educativa es ampliamente reconocida. Sin embargo, existe una controversia muy fuerte, con respecto a los métodos que debe utilizar y el papel que debe jugar la medición en el desarrollo de las evaluaciones.

Por otra parte, el término “assessment” en inglés, para el que, como ya dijimos, no existe un término propio en el idioma español, se reserva en general para referirse a las personas. Puede requerir la administración de pruebas, o simplemente la clasificación y calificación con base en algunos criterios claramente especificados. Las evaluaciones escolares “school assessments” se basan generalmente en una síntesis informal de una amplia variedad de evidencia, aunque también pueden incluir el uso de los resultados de pruebas, o de las respuestas a escalas de actitud y cuestionarios.

3.2 Perspectivas y conclusiones reflejadas en el análisis de algunas experiencias desarrolladas en países con mayor tradición en evaluación

A. La distinción entre los términos del inglés “evaluation” y “assessment”

Esta distinción evidencia una profundización en el campo, que explica el que hoy en día se considere desafortunado el empleo de expresiones como “evaluación del estudiante” con el significado de evaluación antes discutido, debido a que su uso implica un juicio de valor acerca del rendimiento de un estudiante, con base en una comparación, o en relación con los otros estudiantes (Keeves, 1997).

Dicho significado de evaluación se considera inapropiado y poco deseable debido a que, para evaluar los aprendizajes de un estudiante, se puede desarrollar una escala de rendimiento, y sobre todo, porque el

aprendizaje se entiende como una responsabilidad compartida entre el estudiante, la escuela y el hogar. Desde esta perspectiva, se excluye la determinación del aprovechamiento del estudiante, en relación con el de los otros, pues se parte de una concepción educativa que pone el acento en promover el aprendizaje del estudiante y el logro de estándares de desempeño.

B. Las implicaciones extraídas por el Comité para el Análisis de los Fundamentos de la Evaluación a partir del análisis de los avances y resultados de las Ciencias de la Cognición y de la Medición, para el mejoramiento de la “evaluación” de los estudiantes.

Estas implicaciones las presentan en el último reporte¹, en el cual se refieren a la evaluación entendida como “assessment” tanto del aula como en contextos más amplios, con relación a tres propósitos amplios:

- Apoyo al aprendizaje
- Valoración del rendimiento del estudiante
- Evaluación de programas

Al respecto, el Comité reconoce que el *propósito de la evaluación* impone prioridades, y el contexto en que se use impone limitaciones a su diseño. En consecuencia, establece que es indispensable admitir que **uno solo de esos tipos de evaluación, no resulta apropiado para todos**. Plantea además, que con frecuencia se utiliza una sola evaluación para muchos propósitos, pero que generalmente, **entre más propósitos se pretendan alcanzar con una evaluación, mayores condicionantes y ajustes deben hacerse para cada propósito**.

Asimismo, este reporte destaca que si bien los diferentes tipos de evaluación

pueden aparecer como muy diferentes entre sí, todos comparten los mismos principios. Entre tales principios se encuentra el que plantea que la evaluación representa un proceso de razonamiento, a partir de evidencias, por lo que es únicamente una estimación de lo que una persona sabe y puede hacer. Se puede afirmar, por lo tanto, que por naturaleza, la evaluación tiene siempre cierto grado de imprecisión.

Otro principio común es el de que toda evaluación, independientemente de su tipo, descansa en los tres pilares siguientes:

- Un modelo de cómo los estudiantes representan el conocimiento y desarrollan competencia en un dominio específico.
- Tareas y situaciones que permiten observar el desempeño.
- Un método de interpretación para extraer inferencias de las evidencias de desempeño obtenidas de la manera indicada.

De este principio se deriva el hecho de que, en una amplia variedad de prácticas evaluativas, subyace un modelo muy restringido y poco actualizado respecto al aprendizaje y al desarrollo de competencias.

C. Fortalezas y debilidades de los modelos utilizados en Inglaterra para analizar el desempeño del estudiante y para valorar el currículum

Lesley Saunders (2000), de la Fundación Nacional para la Investigación Educativa en Gales, Inglaterra, presenta un análisis de los modelos utilizados para determinar el desempeño del estudiante y valorar el currículum, con base en la experiencia inglesa, y plantea las fortalezas y debilidades de cada uno. Para el análisis, considera los siguientes modelos:

- c.1 Los modelos que se desarrollan con base en muestras de estudiantes.
- c.2 Los modelos universales que requieren una representación sistemática de cómo es el desempeño y progreso de todos y cada uno de los estudiantes.
- c.3 Los modelos que se dirigen a valorar el desempeño institucional.

La descripción de cada modelo, la forma en que se ha empleado en Inglaterra, y las principales fortalezas y debilidades potenciales que se le han documentado, se resume a continuación:

c.1 Los modelos que se desarrollan con base en muestras de estudiantes

Responden al interés de obtener una perspectiva general del aprovechamiento en el nivel nacional, en áreas clave del currículum, por lo que se utiliza **una muestra de la población**, a fin de limitar los costos en términos de recursos y de efectos reactivos en el currículum. Entre sus ventajas, Saunders señala la reducción al mínimo de las complicaciones de la evaluación, en tanto que permite una cobertura amplia del desempeño de los estudiantes en distintas edades y tiempos, ofreciendo una descripción detallada del desempeño curricular en áreas claves, y la identificación de factores escolares y estudiantiles asociados a éste. Requiere un muestreo apropiado y relaciones adecuadas con las autoridades educativas locales y con las instituciones escolares. Entre sus debilidades potenciales anota cuestionamientos relativos a la validez y confiabilidad, y menciona que este modelo generó un debate internacional acerca del denominado "Rash model" y numerosas críticas.

c.2 Los modelos universales

Requieren una representación sistemática de cómo es el desempeño y progreso

de todos y cada uno de los estudiantes, responden al interés de contar con una representación sistemática de los logros y progresos de todos y cada uno de los estudiantes.

La forma en que este modelo se ha venido desarrollando en Inglaterra desde 1988 es mediante el establecimiento, dentro del currículum nacional, de qué es lo que se espera que los estudiantes sean capaces de hacer, y luego, recolectando evidencias para determinar si son o no capaces de hacerlo. Se hicieron ajustes en las metas de desempeño especificadas en el currículum nacional, a efectos de evaluar y reportar el desempeño de los estudiantes a los 7, 11, 14 y 16 años, edades que corresponden con la finalización de las etapas en que se divide el currículum nacional. La evaluación se ha desarrollado mediante una **combinación de pruebas estandarizadas externas y la evaluación de los profesores de las disciplinas centrales.** Su desarrollo requiere una tarea masiva, una gran inversión de recursos financieros y de otros tipos.

Este tipo de modelo requiere una amplia gama de ajustes y afinamientos, como por ejemplo, entre la representación de estándares ideales y la práctica real que se concreta en las aulas; entre precisión y coherencia; entre amplitud y flexibilidad, o entre confiabilidad y congruencia con el aprendizaje real en el aula. La cantidad de reportes y debates acerca de estos aspectos es muy grande y extensa.

Además, las propuestas para este modelo incluyeron su utilización para valorar y comparar el desempeño de las escuelas. Fue utilizado por el gobierno para que las escuelas diesen cuenta públicamente de sus desempeños. Se argumentó al respecto, que los padres utilizarían sus resultados para seleccionar las escuelas y que la puesta en práctica de relaciones de oferta y demanda educativa, contribuiría a elevar los estándares.

Al identificar las fortalezas y debilidades de este modelo, Saunders plantea la dificultad que se deriva de la conjugación de las muy diversas agendas a las que responde este modelo. Sin embargo, reconoce que este extenso programa de desarrollo de pruebas nacionales y de evaluación ha involucrado activamente a muchos educadores, y a otros profesionales, en la creación de una comprensión colectiva de los propósitos y significados del "assessment", independientemente de las limitaciones que se le atribuyen al sistema, tal y como existe en la actualidad. También ha logrado que en la actualidad, difícilmente se presenten argumentos convincentes para oponerse a la documentación y análisis (assessment) del aprendizaje de todos los estudiantes, y a la identificación de cualquier resultado inferior a lo esperado.

Con respecto a sus debilidades potenciales, Saunders señala que además de los cuestionamientos a la validez y confiabilidad que se le apuntan a otros modelos, existen otras debilidades asociadas a éste. Dado que el sistema se aplica a toda la población o universo, y no emplea algún tipo de muestreo, y puesto que los resultados se utilizan también para comparar el desempeño de las escuelas, por lo que cada vez con mayor frecuencia se han venido utilizando también para comparar el de los educadores, existe un riesgo real de que se produzca la "**enseñanza para la prueba**". Es decir, que en la realidad de la mayoría de los salones de clase, el rendimiento en las pruebas se convierte en la directriz curricular y de la enseñanza, en lugar de como conceptualmente se esperaría. Se ha generado una proliferación de pruebas para los grupos de los ciclos en que no corresponde su aplicación, que responde al interés de los educadores por predecir el desempeño futuro a partir del actual. El riesgo o temor mayor, es que como resultado de tal situación, se produzca un **empobrecimiento del currículum**, aunque esto no ha sido concluyentemente demostrado. Peor aún,

algunos críticos argumentan que, puesto que este sistema pretende responder a **múltiples propósitos**, algunos de los cuales han resultado contradictorios entre sí, finalmente **no responde a ninguno de ellos**.

c.3 Los modelos que se dirigen a valorar el desempeño institucional

Estos modelos han adquirido alguna notoriedad en Inglaterra; la que puede atribuirse, al menos parcialmente, a la publicidad que se le ha dado a los resultados de las instituciones escolares, presentados en orden jerárquico de desempeño –destacando y reprimiendo–, en la prensa local y nacional. Esta iniciativa cuenta con apoyo por parte de diversos grupos, además del de las agencias gubernamentales y de los responsables de las políticas. Saunders respalda esta apreciación, haciendo referencia a un artículo en el que un director curricular (Downes, 1998) argumenta que “la apertura de las escuelas al interés público constituye un desarrollo bienvenido que ha sido ignorado durante mucho tiempo”.

Las propuestas originales que pretendían utilizar el desempeño de los estudiantes para valorar el de las escuelas, fueron criticadas por Cuttanse y Goldstein (1988) argumentando que este sistema podría resultar injusto tanto para las escuelas como potenciales proveedoras, como para los padres de familia como seleccionadores potenciales de calidad educativa. En su justificación, estos autores señalan que si el sistema, apoyado en el logro de sus estudiantes, ha realizado comparaciones entre escuelas, en circunstancias donde han existido variaciones de carácter social y/o académico, respecto a las condiciones de ingreso de los estudiantes –y es posible asumir que normalmente encontraremos tales variaciones en el mundo real– esta situación *encubrirá el grado en que las distintas escuelas contribuyeron al progreso realizado por los estudiantes*, distinguiéndolo de los

estándares alcanzados por los estudiantes. Este razonamiento, expuesto en muchos artículos y reportes, ha constituido la objeción esencial al uso que hizo el gobierno de las tablas de rendimiento basadas en resultados no procesados.

Se han dado dos tipos de respuesta a las tablas de ordenamiento jerárquico del desempeño, por parte de la profesión educativa. Una es demandar su abolición completa, –si bien los proponentes de esta medida están seguramente dispuestos a indicar lo que substituirían a efectos de satisfacer los requerimientos legítimos para la rendición de cuentas–.

La otra respuesta es una propuesta para desarrollar mediciones del “valor agregado” en el desempeño; es decir, mediciones que tomen en consideración la relevancia estadística de aquellas variables propias de los antecedentes de los estudiantes, que como bien se sabe, están correlacionadas con el rendimiento. Esto se ha venido realizando en el último sistema de indicadores de rendimiento, y se conoce como “el Paquete Automático”. Sin embargo, este esfuerzo está repleto de dificultades técnicas y operacionales, especialmente en relación con la construcción de indicadores que resulten fácilmente accesibles para una audiencia que carece de conocimientos técnicos, la cual constituye la población para la cual se diseñaron este tipo de tablas: los padres de familia y el público. Una revisión de los desarrollos y aspectos relativos al “valor agregado”, puede consultarse en Foxman (1997) y Saunders (1999).

Entre las fortalezas que se reconocen a este tipo de tablas de rendimiento, es que estas han abierto a las escuelas, como nunca antes, al escrutinio público, y que gran parte del debate que han generado, tanto dentro como fuera de la profesión, se ha desarrollado con buenos fundamentos y ha abordado tanto principios fundamentales como tecnicismos, trivialidades y “defensa de territorios”. Una consecuencia de ello es que el personal escolar tiene una

mejor comprensión de los usos y límites de los datos estadísticos en un nivel jamás imaginado en la década anterior.

Entre las debilidades potenciales que se señalan a este modelo, se menciona el riesgo de que el énfasis en la medición del rendimiento y del análisis del valor agregado, llegue a constituirse en un “*distractor*” para el personal escolar, en una época en que tiene aún pendiente la atención de una gigantesca agenda educativa. Otro de los riesgos que se discuten, más fundamental aún, es el de haber *invertido considerables esfuerzos tanto en la arena política, de investigación, como en la de los profesionales en servicio, en una actividad en que por un lado, podría terminar sin que se pueda lograr reflejarle al público la complejidad del rendimiento escolar, y por otro, que finalmente no ofrezca una garantía para el mejoramiento de los estándares de todas las escuelas.*

D. Pronunciamiento de la Asociación Americana de Evaluación sobre la utilización de pruebas estandarizadas en el nivel escolar

Después de dos años de revisión, exploración y análisis, la Asociación Americana de Evaluación (AEA) aprobó y publicó en febrero de este año (2002), un Pronunciamiento sobre la Utilización de Pruebas Estandarizadas en el Nivel Escolar –equivalente a la Educación General Básica y Ciclo Diversificado-. El desarrollo de este Pronunciamiento contó con financiamiento parcial de la Asociación Americana para la Ciencia. El trabajo fue realizado por un equipo integrado con el propósito de explorar la necesidad de establecer un pronunciamiento, y prepararlo en caso de que lo fuese.

El análisis se inició con una sesión realizada durante la Reunión Anual de la AEA en el año 2000, en la que se hicieron varias presentaciones acerca de diversos aspectos relacionados con este tipo de

pruebas, que fueron luego discutidas por los asistentes. Durante el año siguiente, el equipo responsable debatió, formuló y nuevamente discutió y reformuló su pronunciamiento. Durante la Reunión Anual de la AEA en el año 2001, realizó nuevamente una sesión con el propósito de realimentar el documento, y paralelamente entregó a la directiva de la AEA un borrador del trabajo, para obtener sus criterios y fortalecerlo. Una vez finalizada la reunión, solicitaron opiniones y comentarios al documento por medio de su lista temática electrónica EVALTALK. Con todos esos insumos, el equipo responsable revisó una vez más el borrador del documento, y lo presentó para su aprobación a la directiva de la AEA.

La AEA es una asociación profesional internacional formada por evaluadores, cuya misión es mejorar las prácticas y métodos de evaluación y promover sus contribuciones para la generación de teoría y conocimiento acerca de la acción humana efectiva.

El Pronunciamiento de la AEA, el cual puede consultarse en el sitio web: <http://www.eval.org>, establece como planteamiento central, alrededor del cual desarrolla su argumentación, el siguiente:

Las pruebas estandarizadas para la toma de decisiones críticas conducen a atender inadecuadamente a todos los estudiantes, especialmente a los más necesitados y vulnerables, violando así el principio de “no causar daño”. La Asociación Americana de Evaluación se opone al uso de pruebas como el único o como criterio principal para tomar decisiones que tengan serias consecuencias negativas para estudiantes, educadores, escuelas y colegios. La AEA respalda aquellos sistemas de documentación del aprendizaje “assessment” y rendimiento de cuentas que contribuyen a la educación.

El Pronunciamiento surge de la preocupación de que durante los últimos años, se ha presenciado en los Estados Unidos, una confianza creciente en el uso de pruebas estandarizadas en las escuelas, sin que

se cuente con una validación completa a través de todo el país. Esa confianza se manifiesta en su utilización para tomar decisiones críticas acerca de los estudiantes, educadores, y las escuelas. Dicho aumento de la aplicación de pruebas estandarizadas en el nivel escolar, lo atribuye a la necesidad de contar con información sólida, que contribuya a que los diseñadores de las políticas educativas, definan políticas y prácticas que aseguren el éxito académico de todos los estudiantes.

Sin embargo, el análisis realizado por la AEA a la evidencia acumulada durante las dos últimas décadas, indica que la utilización de pruebas estandarizadas **no conduce a mejorar las políticas y prácticas educativas**. Existe evidencia de que a pesar de que tales pruebas ocasionalmente contribuyen al mejoramiento de las condiciones de enseñanza y aprendizaje, **generalmente provocan consecuencias educativamente injustas y prácticas inapropiadas**. Los efectos más preocupantes son el *incremento de la deserción, la desprofesionalización de los educadores y administradores educativos, la pérdida de la integridad curricular, la creciente falta de sensibilidad cultural, y la desproporcionada asignación de recursos en los programas de pruebas*, en vez de asegurar la calidad de los educadores y de generar programas educativos de alta calidad.

Si bien en el pronunciamiento plantea la necesidad de más estudios acerca de los efectos perjudiciales de la aplicación de pruebas estandarizadas, la AEA considera que existe evidencia lo suficientemente convincente, como para no apoyar la continuación de esa práctica.

Al respecto, inició un esfuerzo, aún en desarrollo, de generar una bibliografía comentada que constituya un recurso para los investigadores interesados en el uso de pruebas en el nivel escolar. Dicha bibliografía, en la que actualmente se incluyen alrededor de 200 referencias, puede ser consultada en la siguiente dirección

<http://www.eval.org/hstbib.htm>. Asimismo, es posible revisar esa fuente, según las diversas categorías utilizadas para su clasificación. Por ejemplo, si el interés es consultar aquellos trabajos relacionados con el desarrollo y aplicación de pruebas, se pueden revisar directamente los trabajos incluidos en esa categoría, la cual es una de las que se emplean en la clasificación, en la siguiente dirección <http://www.eval.org/hst.test.htm>.

Con el argumento de que la simplista aplicación de una sola prueba o baterías de pruebas para tomar decisiones respecto a niveles, la aprobación de personas o grupos, cuando la escolaridad contemporánea enfrenta serias limitaciones, la AEA afirma que con ello se impide, en vez de promover, el aprendizaje escolar.

Por otro lado, las comparaciones de escuelas y grupos con base en los resultados de las pruebas, propician una enseñanza para la prueba, cuyas características no constituyen un mejoramiento de la enseñanza y el aprendizaje. En su pronunciamiento, la AEA asegura que a pesar de que se han venido empleando por más de dos décadas, las pruebas impuestas por los diversos estados –en los Estados Unidos– no han mejorado la calidad de las escuelas; ni han disminuido las disparidades en el rendimiento académico con respecto a género, etnia, o nivel socioeconómico; tampoco han contribuido al avance del país en términos morales, sociales o económicos.

En consecuencia, la AEA respalda y defiende el rendimiento de cuentas, pero no fundamentado en pruebas. Por lo tanto, se une con este pronunciamiento, a muchas otras asociaciones profesionales en la oposición al uso de pruebas para la toma de decisiones críticas.

Entre las violaciones a los principios establecidos tanto por la AEA y por otros estándares profesionales, que evidencian la apreciación de que el impacto de la utilización de pruebas se ha constituido en

una práctica que causa más daños que beneficios, se citan las siguientes:

- Su empleo sin una meta-evaluación independiente y confiable.
 - Su aplicación para tomar decisiones claves –como aprobación– sin disponer de suficiente evidencia de su validez, y sin haber sometido la documentación técnica al escrutinio público.
 - Exhiben deficiencias tanto en lo que respecta a su adecuación técnica como en la precisión en la asignación y reporte de puntajes.
 - Promueven un control centralizado acerca de lo que debe ser enseñado y de cómo debe hacerse, que no resulta conveniente.
 - Dirigen la oferta educativa hacia la satisfacción de unos estándares monolíticos, estrechos, definidos en función de los criterios centralizados establecidos para las pruebas, en lugar de atender las diversas necesidades de los estudiantes de las escuelas locales.
 - Inducen concepciones estrechas acerca de la enseñanza y la educación en las escuelas, las cuales privan a los estudiantes de su historia, perspectiva cultural, experiencia personal, y de la naturaleza interdisciplinaria de las disciplinas.
 - Reducen el currículum a las materias examinadas, entre las que usualmente se incluyen la lectura, la escritura, y las matemáticas, dejando por fuera las disciplinas que no se evalúan, entre las que generalmente se encuentran las artes, la educación física y algunas ramas de las ciencias.
 - Demandan una cantidad desproporcionada del tiempo de educadores y estudiantes que se reduce a otros objetivos y actividades. Se cita por ejemplo, que se ha llegado a dedicar hasta un 30% del año escolar en la preparación para estas pruebas.
 - Asumen que todos los estudiantes, incluyendo aquellos que tienen necesidades especiales, o que se encuentran en proceso de aprender el idioma, aprenden de la misma forma, al mismo ritmo, y que todos están en capacidad de demostrar sus logros por medio de pruebas estandarizadas.
 - Impulsan, directa o indirectamente, a que los estudiantes que podrían perder la prueba no se presenten en ellas, o que deserten definitivamente de la escuela.
 - En gran medida, representan una medición del ingreso y de las condiciones familiares, así como del grupo cultural, perpetuando discriminaciones sociales.
 - Contribuyen a generar una atmósfera de temor, competición separatista –entre escuelas, maestros y estudiantes–, histeria, que es contraria a la enseñanza y el aprendizaje.²
- Por otra parte, en su pronunciamiento, la AEA establece un conjunto de expectativas para el mejoramiento de la prácticas de evaluación. Para ello, establece un conjunto de principios, de los cuales se extrajeron los siguientes:
- Las decisiones evaluativas importantes deben hacerse con base en criterios múltiples y en muchas mediciones de alta calidad, validadas para usos específicos.

- Los especialistas involucrados en los programas de pruebas estandarizadas deben de tener en consideración no sólo aspectos técnicos y teóricos, sino además aquellas consecuencias como el bienestar de los estudiantes, educadores, escuelas, y la sociedad en general.
 - La administración estatal y local de la educación debe considerar una amplia gama de perspectivas con respecto a lo que conviene más a los estudiantes, las escuelas y la sociedad.
 - Los diseñadores de las políticas educativas y los educadores en ejercicio, deben emplear sistemas de documentación y análisis de los aprendizajes y de rendición de cuentas bien concebidos, que incluyan mediciones múltiples, y realizar un esfuerzo continuo por obtener mejores representaciones respecto a lo que se enseña y se logra aprender.
 - Las decisiones críticas –como la promoción o retención de un nivel a otro, graduación, certificación, clasificación, reconocimientos monetarios y sanciones– respecto a los estudiantes, educadores, y escuelas, no deberían basarse en una sola prueba o batería de pruebas, independientemente de cuántas veces pueda tomarse.
 - Tanto el conocimiento y la experiencia de los educadores profesionales como aquellas pruebas estandarizadas completamente validadas, son importantes para la toma de decisiones educativas.
 - Existe una urgente necesidad de iniciar y financiar evaluaciones externas de los programas de pruebas en todos los estados y sistemas escolares en donde se apliques políticas de este tipo.
 - Es indispensable que los resultados de tales evaluaciones externas, acerca de las consecuencias, tanto positivas como negativas, de los programas de pruebas, se entreguen a los diseñadores de las políticas educativas, padres de familia, educadores, estudiantes, y el público en general.³
- Finalmente, el pronunciamiento de la AEA enfatiza que el problema más serio de las pruebas estandarizadas, es su insistencia en evaluar la educación en una forma estrecha. Su práctica responde a un esfuerzo por tratar la enseñanza y aprendizaje de una forma justa y simple, en un mundo donde la educación es increíblemente compleja, caracterizada por la inequidad en la distribución de oportunidades. Si se pretende estandarizar la educación, deben identificarse los retos generados desde diversas perspectivas, que incluyan los costos y beneficios para los estudiantes de nuestras escuelas.
- El pronunciamiento concluye estableciendo que la educación requiere decisiones acerca de cómo los estudiantes, educadores, y las escuelas deberán sostenerse, mejorarse, y promoverse, pero que los programas de pruebas estandarizadas simplifican excesivamente las decisiones a tomar. Por lo tanto, la AEA se compromete a atender el principio de “no causar daño”, que requiere demandar que se examinen las consecuencias, en situaciones reales, de todos aquellos que se vean afectados, y no únicamente desde la perspectiva de las autoridades. Afirma además que las actuales políticas y prácticas de toma de decisiones con base en pruebas estandarizadas, no cumplen con ofrecer los mecanismos de revisión, meta-evaluación, y validación que demandan los estándares profesionales de la AEA.

3.3 Revisión de los principios y conceptos presentes en los estándares propuestos para la evaluación de la Educación Matemática en el nivel escolar por el Consejo Nacional de Educadores Matemáticos de Estados Unidos

El Consejo Nacional de Educadores Matemáticos de los Estados Unidos –National Council of Teachers of Mathematics (NCTM)– la asociación profesional de educadores matemáticos más grande en ese país, dedicada al mejoramiento de la enseñanza y el aprendizaje de las matemáticas, ha adquirido la responsabilidad de asumir el liderazgo nacional en la educación matemática.

Para cumplir con esta responsabilidad, esta organización ha desarrollado un conjunto de estándares para las matemáticas escolares que se refieren a los contenidos, la enseñanza, y la documentación para el análisis de los aprendizajes “assessment”. Dichos estándares constituyen orientaciones para guiar el planeamiento, implementación y evaluación de programas de matemáticas de alta calidad, desde la educación preescolar hasta el ciclo diversificado.

Los conjuntos de estándares que ha generado la NCTM en la última década, son consistentes con los resultados recientes de investigación, pero a la vez, plantean una nueva agenda de investigación, que la NCTM ha venido impulsando.

Los tipos de estándares, así como la secuencia en que la NCTM los ha publicado son los siguientes:

1. Los *Estándares para el Curriculum y la Evaluación de las Matemáticas Escolares* fueron en 1989.
2. Los *Estándares para la Enseñanza Profesional de las Matemáticas Escolares* en 1991.

3. Los *Estándares para la Documentación y Análisis de los Aprendizajes de las Matemáticas Escolares* –“Assessment”– en 1995.
4. Y los *Principios y Estándares para las Matemáticas Escolares* en el año 2000.

Estos estándares han tenido un importante impacto en el nivel internacional, el cual se ha visto reflejado en las discusiones y análisis que han generado en las publicaciones especializadas provenientes de los diversos países del mundo.

Los siguientes, son algunos de sus principios orientadores:

- Todos los estudiantes merecen un programa de matemáticas excelente que los rete a lograr un rendimiento que corresponda con el alto nivel que se requiere para una participación ciudadana responsable y productiva.
- Todos los estudiantes deben contar con educadores calificados, con sólidos conocimientos tanto matemáticos como de las formas en que los estudiantes aprenden matemáticas, y que cuenten con altas expectativas para sí mismos y sus estudiantes.
- El aprendizaje de las matemáticas se maximiza cuando los educadores se centran en el pensamiento y razonamiento matemático. El razonamiento formal y la prueba matemática deben integrarse gradualmente en el programa de matemáticas, conforme el estudiante avanza su escolaridad.
- El aprendizaje de las matemáticas se fortalece cuando el contenido se ubica en contextos específicos y se vincula con materias de otras áreas, y cuando como parte del proceso de aprendizaje, los estudiantes reciben múltiples posibilidades de aplicar

las matemáticas de maneras que tengan sentido para ellos.

- La extensión del impacto de la tecnología en prácticamente todos los aspectos de la vida actual, demanda cambios en los contenidos y naturaleza de las matemáticas de los programas escolares. Para ello, los estudiantes deben ser capaces de usar calculadoras y computadoras en la investigación de conceptos matemáticos y en el logro de una mejor comprensión matemática.
- Los estudiantes emplean diversas estrategias y distintos algoritmos para resolver problemas, y los educadores deben reconocer y sacar ventaja de esos procedimientos alternativos para ayudar a los estudiantes a desarrollar una mejor comprensión de las matemáticas.
- La documentación para el análisis de la comprensión matemática “assessment” debe estar directamente relacionada con el contenido enseñando y debe incorporar múltiples fuentes de información, incluyendo pruebas estandarizadas, pruebas cortas, observaciones, tareas de desarrollo, e investigaciones matemáticas.
- El mejoramiento de la enseñanza y aprendizaje de las matemáticas debe estar guiado por el desarrollo de investigación y de documentación para valorar los programas de matemáticas escolares “assessments”–.
- Cambiar los programas de matemáticas de manera que reflejen estas perspectivas, requiere colaboración de esfuerzos y discusiones permanentes entre todos los grupos interesados en este proceso. La NCTM está dispuesta a trabajar con todos

aquellos que tienen interés en el mejoramiento de la educación matemática para todos los estudiantes. Es a través de esos diálogos y esfuerzos cooperativos que podemos mejorar la competencia matemática de los estudiantes en las clases de matemáticas en todo el continente.

La visión de la educación matemática impulsada por la NCTM intenta superar las formas en que se han organizado, enseñado, documentado y analizado el aprendizaje de las matemáticas escolares en el pasado, las cuales subestimaban la capacidad matemática de la mayoría de los estudiantes y perpetuaban mitos dañinos acerca de las habilidades y esfuerzo de los estudiantes.

A pesar de que reconoce que en la actualidad, dentro de una sociedad multicultural, el desarrollo matemático de cada estudiante debe valorarse con frecuencia, considera sin embargo, que pruebas diseñadas para otros propósitos se han utilizado, sin intención, como filtros para negar a ciertos grupos el acceso a estudios más avanzados de matemáticas.

Al respecto, insiste en que los *procedimientos para valorar los aprendizajes*, no deben utilizarse más para negar oportunidades de aprender matemáticas valiosas. Por el contrario, esa valoración de los aprendizajes debe constituirse en una forma de *promover el desarrollo de altas expectativas*. De no hacerse así, se estaría desperdiciando el potencial humano. Es por ello que la NCTM considera que para alcanzar su visión acerca de la educación matemática, debe impulsarse un cambio de todos los aspectos de las matemáticas en conjunto:

Contenidos:

Hacia una rica variedad de tópicos matemáticos y de situaciones problema, apartándose de un trabajo centrado en la aritmética.

Aprendizaje:

Hacia la investigación de problemas, apartándose de la memorización y repetición.

Enseñanza:

Hacia la formulación de preguntas y escucha, apartándose de la exposición y relato.

Documentación y análisis:

Hacia la búsqueda de evidencias de varias fuentes determinadas por los maestros, apartándose del juicio externo basado en una sola prueba.

Expectativas:

Hacia el uso de conceptos y procedimientos para la resolución de problemas, apartándose del dominio de conceptos y procedimientos aislados.

La respuesta entusiasta por parte de muchos educadores y escuelas a dicha propuesta, se ha visto reflejada en los cambios del contenido de los cursos y en las formas de enseñarlos. Sin embargo, estos cambios tuvieron que enfrentarse al dilema de que los resultados de sus esfuerzos para lograr las nuevas metas, no contaban con el respaldo de las formas tradicionales de documentación y análisis de los aprendizajes –“assessment”–, las cuales no resultan consistentes con esas perspectivas.

En consecuencia, se generó la necesidad de establecer los Estándares para la Documentación y Análisis de los Aprendizajes (1995) que se presentan a continuación, los cuales incluyen los valores y objetivos asociados con las formas de documentación y análisis de los aprendizajes que deben lograrse para alcanzar la visión y cambios propuestos.

La perspectiva básica presente en estos estándares es que resulta esencial alejarse

del enfoque predominante de “ordenamiento jerárquico de rendimientos” hacia un enfoque filosóficamente más consistente con la visión de la NCTM de las matemáticas escolares y de la clase de matemáticas. En vez de asumir como propósito de la “evaluación” de los aprendizajes, el ordenamiento jerárquico de los estudiantes, de acuerdo con un aspecto o característica determinada, el enfoque propuesto, y actualmente en desarrollo en muchas escuelas, asume que se pueden establecer expectativas públicas altas, de manera que cada estudiante se esfuerce por alcanzarlas. Que diferentes formas de desempeño pueden y lograrán satisfacer las expectativas acordadas. Y que los maestros y profesores son quienes se encuentran en las mejores condiciones de documentar y analizar los aprendizajes de los estudiantes y pueden ser jueces apropiados y consistentes de la diversidad de desempeños.

Estos estándares para la documentación y análisis de los aprendizajes proponen que:

- Los maestros son las personas que se encuentran en mejor posición para juzgar cómo se desarrolla el progreso de los estudiantes, por lo que deben considerarse los jueces principales de los estudiantes. Sin embargo, dependiendo del propósito, otros jueces, como los mismos estudiantes, pueden valorar su propio progreso.
- La valoración del aprendizaje de los estudiantes debe corresponder con, y ser parte integral del proceso educativo.
- Deben utilizarse múltiples fuentes de información para la valoración de aprendizajes.
- Los métodos de valoración de los aprendizajes deben ser apropiados para el propósito con que serán utilizados.

- Deben valorarse todos los aspectos del conocimiento y conexiones matemáticas.
 - La enseñanza y el currículum deben considerarse equitativamente al juzgar la calidad de un programa.
- 4. Perspectivas de los grupos responsables de la educación Matemática en Costa Rica, respecto al papel que tienen las pruebas en su problemática: resultados del estudio realizado**
- 4.1 Perspectivas ofrecidas por el grupo de formadores de docentes e investigadores en el campo**
- La prueba de bachillerato constituye una de las categorías que se derivan del análisis y discusión de este grupo de profesionales, respecto a los problemas principales que aquejan a la educación matemática en el nivel secundario en nuestro país. La forma en que consideran su impacto en ella, se describe en los siguientes términos:
- *Meta única de la enseñanza:* Estudiantes, director, profesores y padres de familia de los colegios secundarios ven el examen de matemática de bachillerato como la meta de la enseñanza de esta asignatura.
 - *Único criterio de calidad:* La aprobación del examen de matemática de bachillerato o de noveno año, se asume como indicador de que el estudiante tiene conocimiento matemático, de que cuenta con un buen profesor, e incluso el porcentaje de aprobados por institución se utiliza para discriminar la calidad de la respectiva institución educativa. En consecuencia, el trabajo en noveno y undécimo año se centra en la preparación para estas pruebas.
 - *Promueve memorización y asimilación mecánica de algoritmos y no desarrollo del pensamiento matemático y la construcción de significados matemáticos:* Lo que evalúan el examen de bachillerato son procesos memorísticos y la aplicación mecánica de algoritmos. Puesto que los profesores en ejercicio son conscientes de ello, y el resultado de la prueba es el criterio con el que se juzga la calidad del trabajo, a ella se dirigen todos los esfuerzos, y por lo tanto el trabajo docente en matemática se reduce a entrenar a los estudiantes para que evalúen las opciones de respuesta con una calculadora a la par, y a memorizar y aplicar algoritmos mecánicamente o con base en claves a prueba de razonamiento.
 - *Contribuye a reforzar la imagen y predisposición negativa hacia la disciplina:* El bajo rendimiento que se obtiene en la prueba, que la convierte en el principal obstáculo para graduarse de secundaria, contribuye a reforzar la imagen y predisposición negativa hacia la disciplina.
- En relación con las propuestas planteadas por este grupo, para superar las consecuencias negativas de las pruebas, únicamente se incluye la necesidad de buscar alternativas de evaluación externa, la superación de las limitaciones identificadas en las pruebas actuales, de manera que respondan a las perspectivas que se hayan consensuado para orientar la educación matemática en el país.
- 4.2 Perspectivas ofrecidas por el grupo de asesores nacionales y regionales de matemáticas**
- El papel de las pruebas en la problemática de la educación matemática en el nivel secundario fue abordado en la reflexión

y discusión de este grupo como parte de un problema más amplio, como es el de la carencia de mecanismos de coordinación que garanticen una gestión coherente de la política educativa. Al referirse a este problema, los participantes subrayaron que dicha carencia se percibe especialmente en lo relativo a la evaluación.

A continuación se presenta una síntesis de las consideraciones expresadas al respecto:

A pesar de que se reconoce la importancia de una adecuada evaluación externa y la rendición de cuentas, especialmente para la toma de decisiones para el financiamiento de la educación, se percibe falta de claridad y de coherencia en los criterios de evaluación. Existen orientaciones distintas para el trabajo en el aula, para la producción de material didáctico y libros de texto, y para la evaluación, tanto interna como externa. Con respecto a la evaluación, se evidencia descoordinación entre el Departamento de Control de Calidad del Ministerio, responsable de la evaluación externa, y el Departamento de Evaluación de ese Ministerio, encargado de proponer el Reglamento de Evaluación que se utiliza para valorar el trabajo que se desarrolla en los colegios. Se dan además incongruencias respecto a los requerimientos para la evaluación. Como ejemplo se cita el requisito de preparar cuadro de balanceo en las pruebas para la evaluación interna, y la selección de objetivos al azar en las pruebas para la evaluación externa; también se hace referencia a la falta de apoyo logístico que se ofrece a los docentes para las pruebas externas. Otro elemento de la evaluación que fue cuestionado por este grupo, fue la decisión de semestralizar la evaluación, la cual se tomó sin considerar las implicaciones que tiene para los estudiantes, ya que les resta oportunidades de superación.

Entre los aportes ofrecidos por este grupo de profesionales, a la petición de ofrecer alternativas para superar esta problemática, se encuentra la de precisar e integrar coherentemente los diversos criterios existentes. En su opinión, existe un conjunto de criterios nacionales para guiar la educación matemática, pero se encuentran dispersos e implícitos en documentos como los programas de estudio, el complejo didáctico, el reglamento de evaluación, los criterios para la evaluación externa, etc. Por lo tanto, es necesario recogerlos, interpretarlos, integrarlos, analizar su coherencia, y hacerlos explícitos.

Por otra parte, al expresar sus expectativas de cambio en la educación matemática, este grupo señaló la necesidad de realizar una reforma integral, que promueva una nueva cultura del aula de matemática y

que sustituya a los cambios parciales, que luego desaparecen entre los múltiples y muy diversos aspectos que inciden en los problemas actuales. El sistema de evaluación vigente, fue uno de los aspectos identificados, entre otros, como indispensables de reestructurar con dicha reforma.

4.3 Perspectivas ofrecidas por el grupo de profesores de matemáticas en servicio en la educación secundaria

La prueba de bachillerato es también una de las categorías que se derivan del análisis y discusión de los grupos de profesores de matemáticas en servicio, acerca de los problemas principales que aquejan a la educación matemática en el nivel secundario en nuestro país. Su impacto en ella, fue descrito en los siguientes términos:

- Se perciben como un mecanismo para establecer ciertos niveles mínimos de conocimientos, y controlar si se alcanzaron o no.
- Tienen una serie de limitaciones y deficiencias, como la poca coherencia entre las demandas que plantean y los reglamentos de evaluación vigentes. A manera de ejemplo se cita la existencia de pruebas especiales para los alumnos de adecuación curricular, los cuales deben luego realizar la misma prueba de bachillerato que el resto de los estudiantes.
- Se han convertido en el principal criterio orientador del currículum y del quehacer docente, pues todos los esfuerzos que se realizan en las aulas de matemáticas costarricenses se orientan a preparar a los estudiantes a pasar la prueba de matemáticas de bachillerato más que a aprender matemáticas.

Por otra parte, con relación a la problemática sentida por los profesores

participantes respecto a la evaluación, se hizo referencia a la desmotivación de los docentes, ocasionada por la responsabilidad que se les atribuye, por el fracaso de sus estudiantes, y por la culpabilidad que les genera verificar que fallaron sus esfuerzos por “transmitir” los contenidos programáticos. Afirman que en general se tiene una imagen negativa de su labor, la cual es reforzada por la prensa y por la forma en que el M.E.P. le ofrece la información acerca de los resultados de las pruebas, ya que únicamente entrega los resultados por institución, dejando por fuera la descripción de las condiciones de trabajo que se tienen en ellas, y los esfuerzos que realizan los docentes que laboran en muchas limitaciones, las cuales distan mucho de aquellas que generalmente caracterizan a las instituciones con mejores resultados.

5. Retos, posibilidades, y orientaciones principales que demanda la valoración de la calidad de la educación nacional: a manera de conclusión

El análisis de las perspectivas ofrecidas por los representantes de los tres grupos responsables de la educación matemática, consultados acerca de los aspectos que determinan la problemática de la educación matemática en el país, evidencian un conjunto de coincidencias respecto al papel que tienen las pruebas nacionales, entre sí, y especialmente, con los diversos planteamientos que se desprenden del análisis de los modelos de evaluación que se han venido empleando en el nivel internacional.

Las referencias consideradas para enmarcar el análisis, nos permiten identificar similitudes entre el modelo de evaluación externa empleado en nuestro país, y el modelo “universal” inglés. Dicha similitud se sustenta fundamentalmente en los siguientes aspectos:

- Ambos responden al interés de disponer de una representación sistemática de los logros y progresos de todos y cada uno de los estudiantes, por lo que el sistema se aplica a toda la población o universo, y no emplean ningún tipo de muestreo.
- Utilizan una combinación de pruebas estandarizadas externas y la evaluación de los profesores de las disciplinas centrales.
- Su desarrollo requiere una tarea masiva, una gran inversión de recursos financieros y de otros tipos.
- También se utilizan, cada vez con mayor frecuencia, para valorar y comparar el desempeño de las escuelas.
- Requieren del establecimiento de un currículum nacional básico.

Es posible, sin embargo, puntualizar algunas diferencias técnicas, especialmente en lo relativo al referente que se utiliza para realizar las comparaciones necesarias, para emitir los juicios que requiere la evaluación. El sistema inglés utiliza criterios derivados de los objetivos establecidos en el currículum nacional –de ahí que se haga referencia a ese tipo de pruebas como “pruebas con referencia a criterios”–. En nuestro país hemos utilizado tanto las pruebas con referencia a criterios, como con referencia a normas. En estas últimas, el aprovechamiento de los estudiantes se determina con base en comparaciones con los resultados de los demás, perspectiva que actualmente, en el contexto de los países con mayor desarrollo en el campo de la evaluación, se considera inapropiada y poco deseable, puesto que riñe con una concepción educativa cuyo énfasis está en promover el aprendizaje y en el logro de estándares de desempeño.

Las fortalezas de nuestro sistema de evaluación, coinciden también con las del

modelo universal inglés, ya que las perspectivas de los grupos participantes en la consulta cuyos resultados hemos reportado, evidencian que si bien se plantean serias objeciones al sistema de evaluación externa vigente, y particularmente a los efectos del tipo de pruebas que se han venido empleando, se reconoce la importancia del análisis y valoración de los aprendizajes y la necesidad de una rendición de cuentas. Prueba de ello es que los grupos consultados no plantean la eliminación del sistema de evaluación externa, sino de buscar alternativas para superar las limitaciones señaladas, de manera que respondan a las perspectivas consensuadas (formadores de profesores, asesores), para orientar la educación matemática en el país, o que se tomen en consideración las diferencias propias de las condiciones de trabajo, que necesariamente inciden en los resultados que se publican, y la forma en que se presentan estos resultados (profesores en servicio).

En cuanto a las debilidades del modelo, los trabajos consultados las exponen con claridad. Al respecto, se ha hecho referencia a los requerimientos que plantea ese tipo de modelo, en términos de una amplia gama de ajustes y afinamientos; como por ejemplo, entre la representación de estándares ideales, y la práctica real que se concreta en las aulas; entre precisión y coherencia; entre amplitud y flexibilidad, o entre confiabilidad y congruencia con el aprendizaje real en el aula. Se ha dejado constancia de que la cantidad de reportes y debates acerca de estos aspectos es muy grande y extensa. Asimismo, se señaló la dificultad que se deriva de la conjugación de las muy diversas agendas a las que responde este modelo.

Sin embargo, la seriedad de las implicaciones denunciadas por la AEA, y comentadas ampliamente en este trabajo, y sobre todo, la gran similitud de las apreciaciones brindadas por los grupos responsables de la educación matemática en nuestro país, al referirse a su impacto en

la educación matemática, tornan indispensable una valoración de las pruebas nacionales, y sobre todo de la forma en que éstas se han venido empleando para tomar decisiones que permitan determinar:

- 5.1 El peso que se le ha venido asignando a sus resultados, para la toma de decisiones críticas que tienen consecuencias importantes en la vida de los estudiantes.
- 5.2 El aporte concreto que han ofrecido hasta la fecha, en el mejoramiento de la calidad de la educación matemática en nuestro país.
- 5.3 La conveniencia de realizar una evaluación externa al Ministerio de Educación Pública, o meta-evaluación, que dé cuenta de su contribución real para alcanzar los objetivos que justifican su empleo, en el marco de las demandas presupuestarias y financieras que requieren, y de los eventuales efectos no deseados que puedan estar generando, y que se han reportado en otros países.
- 5.4 La identificación de los propósitos a los que se pretende responder con ellas, las prioridades asignadas a cada uno de esos propósitos, y las alternativas disponibles para alcanzarlas, sin incurrir en las consecuencias negativas que se han documentado en otros países.
- 5.5 El estudio y análisis con profundidad de los documentos que se han presentado como evidencia del daño ocasionado por su uso inapropiado.
- 5.6 El desarrollo de estudios que expliquen las perspectivas teóricas que subyacen al modelo utilizado en la actualidad, y las contrasten con los que se desean promover, respecto a:

- La concepción de aprendizaje de las matemáticas.
 - Las formas disponibles para que los estudiantes puedan representar sus conocimientos y desarrollo de competencias matemáticas.
 - Los tipos de tareas y situaciones más apropiadas para ello.
 - Los métodos que resultan más apropiados para inferir o estimar los aprendizajes y competencias desarrollados por los estudiantes.
- 5.7 La necesidad de definir unos principios orientadores de la evaluación, que guíen los esfuerzos que como parte de ella se realizan, en direcciones convergentes con los propósitos educativos a los que responden, a la vez que impidan que requisitos meramente técnicos distorsionen tales propósitos. Derivar de tales principios, un conjunto de estándares éticos y técnicos que garanticen que su utilización no se desvíe de los propósitos que la impulsan.

Notas

1. Knowing what students know (s.f.)
2. Traducción y selección de la autora.
3. Traducción y selección de la autora.

Referencias bibliográficas

American Evaluation Association. High stakes testing of the American

Evaluation Association: Position Statement. <http://www.eval.org>, 2002.

Comité on the Foundations of Assessment. "Knowing what students know". Report.

Contreras, I. "La educación matemática en la escuela secundaria: Mecanismos de resistencia de un modelo agotado". Informe Proyecto de Investigación No. 724-97-347, 2002.

Cuttanse, P. y Goldstein, H. . "A note on national assessment and school

Comparisons" *Journal of Education Policy*. 3 (2) pp. 197-202. 1988.

Downes, "The head's perspective". *Oxford Review of Education*. 24 (1) pp. 25-33. 1998.

Keeves, J.P. "Introduction: Advances in Measurement in Education" *Educational research, methodology, and measurement: An international handbook*. (2 edition). Edited by John P. Keeves. Australia: Pergamon. Pp. 705-712. 1997.

National Council of Teachers of Mathematics (NCTM. *Assessment Standards*. <http://standards.nctm.org/>

Saunders, L. *Value added and measurement of school effectiveness: A critical review*. Slough: NFER. 1999.

_____. *The curriculum and its assessment; Some system models for assessing student performance. The case of England*. Slough: NFER. 2000.

Ileana Contreras Montes de Oca.
Directora Departamento de Educación
Secundaria de la Escuela de
Formación Docente de la
Universidad de Costa Rica.