

ESTABLECIENDO ESTÁNDARES EDUCATIVOS DE DESEMPEÑO EN HONDURAS: EL PRIMER PASO EN CREAR UNA CULTURA DE EVALUACIÓN Y RENDICIÓN DE CUENTAS EN EDUCACIÓN

Jeffrey Lansdale,

Facultad de Postgrado, UNITEC, Tegucigalpa, Honduras

Zarko Vukmirovic

American Institutes for Research, Washington, D.C. U.S.

Mauricio Estrada

Facultad de Postgrado, UNITEC, Tegucigalpa, Honduras

(Noviembre, 2012)

Resumen. Para mejorar la calidad educativa en Honduras, un proyecto educativo y la Secretaría de Educación han desarrollado un sistema de evaluación alineado a los estándares educativos nacionales y el currículo nacional básico. El sistema de evaluación se basa en la definición de estándares de desempeño nacionales, los cuales se utilizan para presentar los resultados de las evaluaciones. Los autores manifiestan que la presentación de resultados de evaluaciones en base al porcentaje de estudiantes en cada nivel de desempeño es de mayor utilidad a las autoridades educativas y otros en el sector educativo que el sistema tradicional de presentar los resultados de evaluaciones en base al porcentaje de respuestas correctas. Ejemplos de las pruebas de fin de grado 2010 en Honduras reflejando ambas maneras de dar a conocer los resultados se presentan en figuras comparativas.

Abstract: In order to improve Honduran educational quality, the ministry of education and an educational project have developed an evaluation system that is in alignment with the national education standards and the national study plans. This evaluation system is based on the national performance standards which are used to represent the evaluation results. The authors of this research paper state that the evaluation results based on the percentage of students in each performance level is more useful to the educational authorities than the traditional evaluation system based on the percentage of correct answers. This paper includes samples of final evaluation tests taken in 2010 in Honduras. It shows the two ways to present the evaluation results and it includes comparative figures.

Palabras Clave: *Honduras, educación pública, currículo nacional, sistemas de evaluación educativa, estándares de contenido, estándares de desempeño, informes de evaluación*

INTRODUCCIÓN

En el año 2002, la cooperación internacional aprobó la propuesta de Honduras de asociarse a la iniciativa global Educación para Todos (EFA, Education for All), consensuada como una de las estrategias para enfrentar la pobreza dentro del marco de las ocho Metas del Milenio acordadas por los integrantes de las Naciones Unidas. A través de EFA, se propone que los países participantes lograrán cobertura de educación primaria universal, con educación de calidad.

Para mejorar la calidad de la educación en Honduras, y lograr las metas de EFA, la Secretaría de Educación (SE) de Honduras desarrolló un nuevo currículo nacional (Diseño Curricular Nacional Básico, DCNB), el cual ha sido llevado a las escuelas de todo el país desde 2004. En apoyo a la implementación del DCNB en las aulas, el Proyecto Mejorando el Impacto al Desempeño Estudiantil en Honduras (MIDEH), (financiado por la Agencia de los Estado Unidos para el Desarrollo Internacional, USAID/Honduras), en conjunto con la SE, definieron los estándares educativos nacionales y desarrollaron las programaciones para acompañarlos (las programaciones contienen los estándares organizados por grado según el mes). Estos documentos fueron distribuidos a los docentes a nivel nacional. Los estándares sirvieron como base para el desarrollo de evaluaciones diagnósticas, formativas mensuales y sumativas de fin de grado. Todos estos materiales están alineados a los libros de texto (Español, Matemáticas, Ciencias Naturales y Ciencias Sociales) los cuales, a su vez, están alineados al currículo. (Se adjunta como apéndice el diagrama del DCNB y los materiales alineados a él).

El nuevo currículo y los materiales de apoyo le brindan a Honduras la oportunidad de realizar una reforma sistémica. Esta oportunidad se ve complementada por el compromiso de los cooperantes internacionales (Alemania, Canadá, España, Estados Unidos, Japón, Suecia, Unión Europea) en apoyar a Honduras a lograr las metas EFA. Mediante este apoyo la SE ha reproducido los estándares y programaciones para todos los docentes del país, así como las pruebas diagnósticas y formativas para cada estudiante del país. También se han destinado recursos nacionales e internacionales para la aplicación de las pruebas sumativas de fin de grado a nivel nacional en los años 2007, 2008 y 2010. Los resultados de estas evaluaciones están siendo utilizados por la SE para diseñar estrategias de capacitación que responden a las necesidades de los docentes según reflejan los resultados de las pruebas. Al implementar esta reforma sistémica de manera exitosa, hay esperanza que Honduras logre una mejora sustancial en la calidad de la educación.

HACIA UNA CULTURA DE EVALUACIÓN Y RENDICIÓN DE CUENTAS

La definición de estándares educativos, y procesos de evaluación alineados a los estándares, representan una oportunidad para la reforma educativa a nivel internacional durante las últimas cuatro décadas. (Glatthorn, 1999; O’Shea, 2005). La reforma en base de estándares (SBR, standards-based reform) enfatiza la importancia del aprendizaje logrado por los estudiantes, en vez de enfocarse en temas como la infraestructura y número de estudiantes por docente. Los estándares definen las expectativas académicas de los estudiantes: los estándares de contenido definen lo que los estudiantes deben saber y saber hacer; los estándares de desempeño establecen los niveles de aprendizaje logrado por los estudiantes. Las evaluaciones alineadas a los estándares sirven para determinar el nivel de dominio de los estándares por los estudiantes.

La reforma en base de estándares es un proceso complejo que requiere la participación de múltiples actores del sector educativo. Por lo tanto, la reforma se ha logrado implementar con diferentes niveles de éxito en países de desarrollo como Canadá y los Estados Unidos, y otros en Asia y Europa. “El cambio exitoso ocurre cuando todos los aspectos del currículo local están alineados a los estándares a través de un sistema coherente y propositivo de procesos y productos.” (Carr y Harris, 2001). En países en vía de desarrollo, la reforma en base de estándares ha sido recién introducida, y mientras se anticipa la mejora de los índices académicos de los estudiantes, aun no se han documentado logros significativos.

En Latinoamérica, según informes del Programa de Promoción de la Reforma Educativa en Latinoamérica y el Caribe (PREAL), la definición de estándares educativos y la implementación de reforma en base de estándares son procesos que asumen mayor importancia a partir del siglo veinte uno. Un informe preparado por PREAL y presentado en 2005 señala que Honduras “cuenta con un nuevo Currículo Nacional Básico y se está definiendo la propuesta de estándares nacionales con un sistema de evaluación de rendimiento académico.” (PREAL, 2005). Sin embargo, el informe, y otro presentado por PREAL en 2011 advierten que las autoridades deben hacer mejor uso de los resultados de las evaluaciones para definir estrategias que impactarán en forma efectiva el rendimiento académico estudiantil.

En Honduras, en 2004, la SE, con el apoyo del Proyecto MIDEH, inició el proceso de la reforma en base de estándares iniciando con la definición de los estándares de contenido. El siguiente paso fue el ordenamiento de los estándares en las programaciones mensuales, donde los estándares fueron organizados y presentados por grado y por mes. Los materiales educativos alineados al DCNB fueron diseñados como herramientas prácticas para los docentes. De hecho,

docentes de todo el país reconocidos como líderes en el sector educativo participaron en talleres organizados para desarrollar los materiales. Los talleres fueron facilitados por expertos nacionales e internacionales en evaluación en base a estándares. Los talleres tuvieron distintos objetivos, tales como definir los estándares educativos, desarrollo de pruebas y redacción de reactivos. Los docentes participantes representaron grados y asignaturas específicas. Al participar en estos talleres, los docentes no solo ayudaron a desarrollar los materiales, también se capacitaron en el proceso. Estos docentes se convirtieron en protagonistas de la implementación de los materiales. Consecuentemente la respuesta de los docentes y directores de centros educativos a la introducción del nuevo currículo y los materiales de apoyo ha sido altamente positiva. (Alas, Hernández & Moncada, 2008).

Esta reforma basada en estándares prosiguió con el desarrollo de un sistema de evaluación tanto interno como externo, con pruebas diagnósticas, formativas y de fin de grado, estandarizadas y alineadas a los estándares de contenido. Hasta ese momento algunos docentes diseñaban sus propias evaluaciones diagnósticas y formativas, con varias implicaciones: la calidad de las pruebas variaba, así como el contenido, lo que hacía imposible comparar resultados incluso entre dos secciones de la misma escuela, mucho menos entre diferentes escuelas. Con las evaluaciones estandarizadas, los docentes se ven obligados a enseñar según lo estipulado en las programaciones, y los resultados de las pruebas son comparables a nivel nacional. Los docentes, directores de escuelas, y autoridades de la SE están en posición de analizar los resultados de las pruebas y tomar decisiones usando datos confiables.

La integración de una reforma programa de evaluación en base a estándares en el sistema de educación nacional crea las bases para una cultura de rendición de cuentas. Los resultados son confiables y los docentes, estudiantes, padres de familia y autoridades educativas están interesados en los resultados y cómo pueden usarse para análisis y toma de decisiones para mejorar la calidad de la educación. El rol primario de la rendición de cuentas en el sector educativo es lograr las metas educativas, a nivel local y nacional. Es por esto que surgió la necesidad de un marco de evaluación unificado a través de los distintos grados y asignaturas, lo que llevó al paso final en la construcción de un sistema de evaluación nacional: la definición de estándares de desempeño.

A pesar de innumerables esfuerzos para empoderar a la Secretaría de Educación y posicionar a sus autoridades como líderes de la reforma curricular y de evaluación, este tipo de innovación ha sido asociado con el Proyecto MIDEH. La USAID y el personal del Proyecto han tomado medidas para minimizar esta atención dirigida al Proyecto, tales como obtener aprobación formal de USAID para reproducir los

documentos de estándares, programaciones y las distintas pruebas sin logos que apunten al Proyecto o al apoyo del gobierno de los Estados Unidos. Existen numerosas razones para justificar la importancia de empoderar a la Secretaría de Educación de Honduras. Una es que los proyectos tienen duración limitada, y MIDEH no es una excepción. La reforma educativa apoyada por el Proyecto debe ser una parte integral de las aulas de clase aún después de finalizar el Proyecto. También es importante notar que, para propósitos de aceptación de la reforma a nivel general, la evaluación en base a estándares ha sido presentada no como parte de la agenda de un proyecto específico si no como una extensión de la reforma sistémica que ha mejorado la calidad de la educación en otros países durante varios años. Ejemplos de liderazgo en esta reforma, como el trabajo sobre evaluación formativa del Dr. James Popham, (Popham, 2006), se usan en Honduras para crear un modelo estratégico y brindar credibilidad internacional al trabajo realizado en el país.

Desarrollo de Estándares de Desempeño para la Medición y Reporte del Aprendizaje de los Estudiantes de Honduras

Los estándares de contenido de Español y Matemáticas han sido definidos y usados ampliamente para guiar el proceso de enseñanza en las aulas a través de las programaciones, a este proceso le siguió la implementación de las evaluaciones formativas y diagnósticas. De acuerdo al estudio de validación externa realizado en 2006-2007, 77% de los profesores de primaria usaban las programaciones para orientar sus planes de clase (Alas et al, 2008). Como siguiente paso fue importante construir un sistema para monitorear los efectos de estas intervenciones en el aula, un sistema de evaluación sumativa para evaluar los aprendizajes de los estudiantes. Este paso requirió el desarrollo e implementación de estándares de desempeño.

Como etapa inicial era necesario explicar el concepto de estándares de desempeño e involucrar al personal de la unidad de evaluación de la SE, Dirección General de Evaluación de la Calidad Educativa (DIGECE). En muchas maneras el reto fue cambiar la cultura de presentar los resultados de las evaluaciones en función de porcentajes. Incluso los indicadores de EFA relacionados al logro estudiantil fueron definidos usando un promedio nacional en porcentaje para las áreas de Español y Matemáticas. Esta es la manera que ha sido comprendida y reportada por las autoridades hondureñas y la comunidad internacional que apoya la iniciativa EFA. Inevitablemente, la sociedad en general ha esperado la presentación de resultados en base a porcentajes debido a que esta es la forma en que tradicionalmente se ha hecho. Sin embargo, al hacer una inspección más cercana a la propuesta de EFA podemos ver que el logro de las metas está definido por el porcentaje de estudiantes

que logran criterios claramente definidos, lo que nos lleva a la conclusión que un criterio basado en logros académicos debe ser medido en función del porcentaje de estudiantes que logran un nivel de desempeño académico definido.

Hay una pregunta muy importante que reta el enfoque tradicional: ¿Qué quiere decir un porcentaje? ¿Es una medida de la dificultad de la prueba o del desempeño del estudiante? ¿Qué significa que un estudiante de noveno grado logró 32% en matemáticas y 63% en español? ¿Qué significa que el promedio de una escuela fue de 56% en cualquiera de las áreas? ¿Realmente sabemos que saben y saben hacer los estudiantes? Dadas estas incógnitas, los estándares basados en porcentaje correcto (por ejemplo 90-100 es excelente, 80-90 es muy bueno, etc.) son vagos y sin sentido preciso. Por lo tanto, necesitamos un marco de referencia más significativo para evaluar el desempeño de los estudiantes. Este inconveniente también se desprende de la perspectiva de la utilidad que tienen los porcentajes para los docentes y autoridades educativas para determinar políticas de forma informada.

Se inició con delinear la diferencia entre estándares de contenido y estándares de desempeño y cómo están relacionados. La siguiente descripción de la diferencia entre ellos se extrajo de una presentación de capacitación y socialización: los estándares de contenido definen qué deben saber y saber hacer los estudiantes, mientras que los estándares de desempeño representan un marco para evaluar cuánto de esos conocimientos y habilidades deben dominar.

Los estándares de desempeño se integraron de manera concurrente en la práctica con el diseño de procedimientos para puntuar las pruebas diagnósticas y formativas orientadas a los docentes. En talleres con el personal de evaluación de la Secretaría de Educación, se establecieron cuatro niveles de desempeño básicos, y se redactaron las definiciones generales para cada uno de dichos niveles. En base a este trabajo, las pruebas y manuales para maestros se diseñaron con instrucciones para que cada docente categorice a los estudiantes en uno de los cuatro niveles en función del número de ítems contestados correctamente. Esto proveyó a los docentes, oportunidades para evaluar a cada estudiante, pero también para medir el desempeño de todos los estudiantes en función de los niveles de desempeño. La presentación de los resultados fue especialmente útil para los directores de centros educativos, quienes pudieron ver cuáles docentes y grados estaban teniendo dificultades para lograr que los estudiantes lleguen a los niveles más altos. Es interesante notar que los docentes y directores de centros respondieron positivamente a la incorporación de este sistema de puntuación en las evaluaciones diagnósticas y formativas, lo cual a su vez sirvió para convencer a algunas de las más escépticas o reticentes autoridades de la SE.

En el desarrollo de los estándares de desempeño para la evaluación sumativa usamos un marco de referencia que ha sido ampliamente investigado y probado en la práctica (Cizek, Bunch & Koons, 2004). Este marco conceptualiza que el proceso de *Definir Estándares de Desempeño (DED)* está compuesto por dos grandes procedimientos: *Definir Niveles de Desempeño*, y *Definir Puntos de Corte*.

Definir Estándares de Desempeño es un procedimiento para conceptualizar los niveles de desempeño que se usarán para evaluar los aprendizajes: se trata de decidir acerca del número y el propósito de los niveles de desempeño, escoger los nombres de los niveles de desempeño (Etiquetas de Niveles de Desempeño) y desarrollar descripciones generales y específicas de los niveles de desempeño (Descriptor de Niveles de Desempeño). Este procedimiento se basa en información extraída de documentación oficial acerca de estándares educativos y el currículo. El procedimiento típicamente se realiza utilizando grupos focales con docentes para recopilar sus opiniones y juicios como los expertos. En Honduras, se le dio gran importancia para formalizar y documentar esta etapa, para poder usar las descripciones de niveles de desempeño para propósitos de interpretación y producción de reportes de resultados.

Definir Puntos de Corte es un procedimiento para establecer los puntos de corte de las pruebas operativas. Estos puntos de corte se utilizan para clasificar a los estudiantes en niveles predefinidos. Hay procedimientos diferentes que se utilizan para determinar los puntos de corte, los cuales se pueden categorizar en dos grandes grupos: procedimientos basados en juicios acerca de los ítems de la prueba; y, procedimientos basados en las personas y su trabajo (Ziky & Perie, 2004). En Honduras se utilizó uno de los métodos más comunes, respaldado por amplia investigación para el proceso de la definición de los puntos de corte en evaluación educativa: el procedimiento de Angoff (Angoff, 1971; Cizek, Bunch & Koons, 2004).

Dentro de cada evento de definición de estándares de desempeño en Honduras (Abril 2008 para 1º, 3º y 6º grados; Diciembre 2008 para 2º, 4º y 5º grados; y Febrero 2011 para 7º, 8º y 9º grados) los dos procedimientos se realizaron de forma consecutiva en cuatro sesiones de ocho horas, las sesiones las realizaron grupos de docentes especialistas en cada área curricular (Español y matemática), por grado. Cada grupo consistió de 15 docentes, apoyados por dos facilitadores por área curricular.

Panelistas fueron docentes representantes de todo el país en función de la ubicación de las escuelas (18 departamentos), rural vs. urbano, género y origen étnica. Los panelistas fueron identificados usando los siguientes criterios:

1. Ser docente del área curricular (Español o matemática) en uno o más de tres grados.
2. Tener al menos 3 años de experiencia como docente en uno de los tres grados.
3. Haber recibido al menos una capacitación durante el servicio (preferiblemente relacionado a evaluación educativa: por ejemplo, talleres de redacción de reactivos, desarrollo de estándares, etc.).
4. Haber implementado los estándares educativos en su trabajo.

Los facilitadores fueron especialistas de contenido de la SE y Proyecto MIDEH, quienes fueron capacitados en el proceso. El Equipo Técnico fue compuesto por especialistas en infotecnología y en psicometría que fueron capacitados para dar apoyo técnico al proceso.

DEFINICIÓN DE NIVELES DE DESEMPEÑO

Los niveles de desempeño se refieren al grado en que los estudiantes dominan los estándares de contenido. La definición de niveles de desempeño es un proceso mediante el cual las autoridades de la SE y los docentes, quienes son considerados los expertos en el aula, son guiados por el psicómetra para definir los nombres que se darán a los niveles de desempeño y para definir *Descriptores de Niveles de Desempeño* generales y específicos. En Honduras, estos talleres se realizaron en Abril del 2008, Diciembre de 2009 y Febrero de 2011, con docentes representando los 18 departamentos, así como representantes de la SE. Durante los talleres se desarrollaron los descriptores de niveles de desempeño para Español y matemática para 1^o-9^o grados. El proceso para definir niveles de desempeño se presenta mediante el siguiente flujograma:

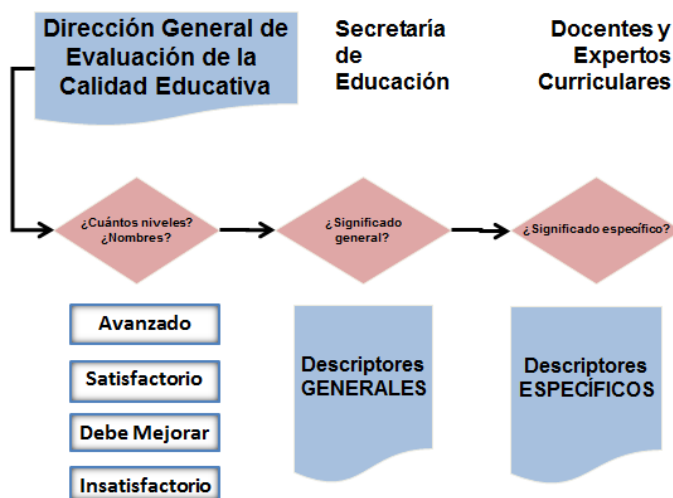


Figura 1: Flujo del proceso para definir niveles de desempeño.

La tarea inicial fue determinar y definir los descriptores generales de niveles de desempeño, con los siguientes resultados, los cuales han sido oficializados por la SE:

1. Insatisfactorio: El/la estudiante demuestra conocimiento muy limitado de los contenidos, y dificultades para resolver problemas sencillos. El/la estudiante demuestra un desempeño muy por debajo del logro mínimo aceptable de los estándares.
2. Debe mejorar: El/la estudiante demuestra un conocimiento parcial de los contenidos y resuelve algunos problemas sencillos. El/la estudiante demuestra un desempeño mínimo aceptable en el logro de los estándares.
3. Satisfactorio: El/la estudiante demuestra buen dominio de los contenidos y capacidad para resolver una variedad de problemas. El/la estudiante demuestra un desempeño aceptable en el logro de los estándares.
4. Avanzado: El/la estudiante demuestra alto dominio de los contenidos y capacidad para resolver problemas con mayor grado de dificultad. El/la estudiante demuestra un desempeño superior en el logro de los estándares.

El siguiente paso fue definir los *Descriptorios Específicos de Niveles de Desempeño*, los cuales determinan lo que los estudiantes deben saber y saber hacer dentro de cada nivel de desempeño para cada grado y para cada área curricular. El proceso típicamente utiliza grupos focales para recolectar las opiniones de todos los participantes, llegando al consenso a través de varias rondas de discusión y análisis.

Las especificaciones que se desarrollan se utilizan en los reportes de resultados y también reflejan lo que se debe enseñar a los estudiantes, de tal manera que se busca identificar recursos e implementar metodologías para ayudar a mejorar el desempeño estudiantil, lo que maximizará los resultados y, por ende, alcanzar la meta de mejorar la calidad de la educación.

Los docentes, padres de familia y el público en general deben poder entender el significado de los puntajes obtenidos por los estudiantes así como lo que deben saber y saber hacer los estudiantes clasificados en cada nivel de desempeño. Pero lo más importante es identificar las áreas en las cuales existe la necesidad de mejorar la calidad de la educación en Honduras.

DEFINICIÓN DE PUNTOS DE CORTE

El procedimiento de Angoff usa ítems de las pruebas para generar los juicios de los docentes expertos en cuanto al desempeño de los llamados “*estudiantes de frontera*” o “*estudiantes con cualidades mínimas*” para cada nivel de desempeño. En el procedimiento de Angoff original los juicios se emiten de forma dicotómica (Sí o No) para contestar a la pregunta “¿*Podría un estudiante con las cualidades mínimas de un nivel de desempeño particular contestar correctamente este ítem?*” Se utilizó el procedimiento de Angoff modificado, el cual solicita juicios que indiquen la probabilidad que los estudiantes de frontera contesten el ítem correctamente. El material requerido para las sesiones comprende los cuadernillos de las pruebas y hojas para respuestas especiales donde los participantes registran sus respuestas.



Figura 2: Concepto de estudiante de frontera y niveles de desempeño

El estudiante de frontera es aquel que sabe y puede hacer todo lo que esta establecido en el nivel de desempeño inferior pero no necesariamente sabe o puede hacer lo que está en el siguiente nivel. Por ejemplo, el estudiante de frontera del nivel satisfactorio sabe y sabe hacer todo lo que se ha definido para el nivel debe mejorar, pero no sabe o sabe hacer todo lo que se especifica para el nivel satisfactorio, este estudiante solo ha logrado los conocimientos y habilidades mínimas para poder ser clasificado en ese nivel. La siguiente figura ilustra este concepto:

El procedimiento de Angoff típicamente se desarrolla en tres rondas. Cada ronda está diseñada para promover el consenso entre los participantes, pero no es requisito llegar al consenso para definir los estándares. Después de cada ronda, se provee retroalimentación para ilustrar el *grado de acuerdo entre los panelistas* y los *datos de impacto* que muestran como los puntos de corte derivados de los juicios influyen en el porcentaje de estudiantes clasificados en cada nivel de desempeño. A los panelistas se les entrega información relacionada con sus propios puntos de corte y los del grupo para evaluar el grado de acuerdo de sus juicios y los del grupo, también se les presentan los datos de impacto para discutir el porcentaje de estudiantes clasificados en cada punto de corte para contestar la siguiente pregunta: “Según su experiencia en el aula, ¿qué tanto refleja esta información lo que realmente sucede en el aula?” Para explicar con más detalle, los pasos usados en Honduras para implementar el método de Angoff fueron:

1. Desarrollar el concepto de *estudiante frontera*. El método requiere que los panelistas discutan acerca de los conocimientos y habilidades que marcan las fronteras entre los niveles de desempeño.
2. Los panelistas emiten juicios para cada ítem en la prueba basado en la experiencia y conocimiento que tienen en la asignatura. Se les instruye para que evalúen cada ítem y contesten la siguiente pregunta para cada nivel de desempeño (excepto “insatisfactorio”).

¿Qué porcentaje de estudiantes de frontera para el nivel de desempeño “**debe mejorar**” (o “**satisfactorio**”, o “**avanzado**”) debería contestar correctamente esta pregunta?

3. Los juicios de los panelistas son analizados y la *retroalimentación* se genera para ser evaluada y discutida después de cada una de las tres rondas. La retroalimentación incluye datos del grado de acuerdo (muestra el grado de acuerdo que tuvieron los panelistas en sus juicios) y datos de impacto (porcentaje de estudiantes que se clasificarían en cada nivel de desempeño si los puntos de cortes propuestos fueran aceptados, total y separado para hombres y mujeres).

Al concluir la tercera ronda, los resultados se presentan y se brinda una última oportunidad a los panelistas para discutir los puntos de corte. Se hace una moderación global para recomendar los puntos de corte finales. Estas recomendaciones finales se deben revisar y aprobar por la SE, para que puedan ser usadas para generar los reportes de resultados anuales.

Los puntos de corte, así como los datos de impacto, se basan en las pruebas operativas de la SE/MIDEH de 2007 para 1º, 3º y 6º grados, en las pruebas de 2008 para 2º, 4º y 5º grados, y en las pruebas 2010 para los grados 7º, 8º y 9º. Es importante hacer notar que los puntos de corte establecidos para una prueba no pueden ser aplicados directamente a una nueva prueba en años subsiguientes, requieren la implementación del proceso de *equiparación* para permitir la comparabilidad entre distintas formas de pruebas. Por ende, los reportes producidos por la SE/MIDEH en el año 2010 para los grados 1º a 6º se basan en *puntos de corte equiparados*, para asegurar que los mismos criterios son usados para clasificar a los estudiantes en niveles de desempeño sin importar la forma de la prueba que se tome. Por esta razón, además de los estándares de desempeño, la equiparación también es parte requerida del sistema nacional de evaluación cuyo fin es monitorear el progreso educativo a través de varios años.

RESULTADOS DE LAS PRUEBAS FIN DE GRADO 2010 DE LA SE/MIDEH EXPRESADAS EN PORCENTAJE CORRECTO Y NIVELES DE DESEMPEÑO

Los resultados de la definición de estándares de desempeño fueron usados como marco para reportar el desempeño estudiantil a nivel nacional, regional y escolar. Ejemplos de la información que ha sido reportada a nivel nacional y en cada departamento basada en los resultados de las pruebas sumativas de fin de grado aplicadas por la SE/MIDEH a una muestra de escuelas a nivel nacional se presentan en el Apéndice 1. Los mismos resultados se presentan en porcentaje correcto así como porcentaje de estudiantes en cada nivel de desempeño. Los dos diferentes métodos se presentan en la misma página para permitir comparar y resaltar la importancia de los reportes por nivel de desempeño para el análisis y toma de decisiones acertada.

Por ejemplo, según los resultados en porcentaje correcto de Español, uno puede concluir que los resultados no están tan malos. Pero, al revisar los resultados por nivel de desempeño, podemos identificar los problemas que existen en el tercer grado, y a la necesidad de prestar especial atención a mejorar la enseñanza de la lectoescritura durante los primeros años de la escuela primaria.

En matemáticas la situación es mucho más crítica. Al revisar las puntuaciones del porcentaje correcto, uno podría concluir que los estudiantes comienzan muy bien el primer grado. Sin embargo, el reporte por niveles de desempeño claramente muestra que incluso en primer grado, el 34% de los estudiantes están en el nivel insatisfactorio y debe mejorar, y es en este punto que los problemas comienzan y se acentúan en los siguientes años.

CONCLUSIONES

En Honduras, el alineamiento que existe entre el currículo nacional, libros de texto y otros materiales educativos ha hecho posible el desarrollo e implementación de un sistema de evaluación basado en estándares de clase mundial. Los materiales de evaluación sirven para medir el nivel académico estudiantil al inicio del año escolar (pruebas diagnósticas), mensualmente (pruebas formativas) y al finalizar el año escolar (pruebas sumativas).

Se han definido los estándares de desempeño de forma conceptual (descriptores de niveles de desempeño) y de forma operativa (puntos de corte); se han presentado reportes de los resultados de las pruebas a nivel nacional, regional y escolar. Estos resultados se presentaron en función de la cantidad de estudiantes en cada nivel de desempeño. Se observó cierto nivel de resistencia respecto al cambio de reportar porcentaje correcto, pero con actividades de promoción en los medios y mediante talleres desarrollados a nivel nacional existe una buena perspectiva que el sistema será aceptado e implementado.

Algunos comentarios y observaciones adicionales acerca del contexto Hondureño son:

1. Los cambios curriculares propuestos a través de la implementación de la evaluación en base a estándares están en consonancia con la política de reforma educativa y el plan de evaluación de la SE descrita en el *Sistema Nacional para la Evaluación de la Calidad Educativa* (SINECE).
2. La incorporación de los niveles de desempeño se ha logrado en las tres evaluaciones: diagnóstica, formativa y sumativa. Esto asegura coherencia en el sistema y la capacidad de producir reportes prácticos y útiles que han sido aceptados por los docentes y directores de centros educativos, mientras que la respuesta del gobierno ha sido de apoyo pero no tan entusiasta.

3. El empuje de la implementación de esta reforma sistémica viene de las aulas, donde docentes y directores se han comprometido a implementar el sistema con impacto cuantificable respecto al logro y aprendizaje de los estudiantes.
4. Los retos más significativos en la implementación exitosa han sido con las autoridades a nivel central, por distintas razones: cambios frecuentes de Ministros y de sus equipos de trabajo de confianza (7 distintos ministros entre 2005-2012); las distracciones y huelgas de los colegios magisteriales; y, el constante cambio de personal técnico debido a la influencia de los políticos.
5. El tiempo de implementación de esta reforma política, para coincidir con la iniciativa de EFA, ha tenido implicaciones positivas importantes. Los fondos proveídos por los cooperantes internacionales fueron clave para la reproducción y distribución de los materiales y capacitaciones a los docentes, estudiantes, padres de familia y autoridades a nivel nacional. Este apoyo ha sido vital para la aplicación de las pruebas sumativas de fin de grado en 2007, 2008, y 2010.
6. El gobierno de Honduras, a través de la SE, ha invertido mucho en la capacitación de los docentes. Estudios externos han demostrado que estas capacitaciones han tenido un impacto limitado en el desempeño estudiantil. Al presentar los resultados de las pruebas por nivel de desempeño a nivel departamental, las estrategias de capacitación se descentralizarán para que los directores de centros educativos y los docentes no solo participen en el diseño y facilitación de las capacitaciones según sus necesidades específicas, también serán responsables de transferir los conocimientos adquiridos a través de la capacitación mediante una enseñanza más efectiva en el aula.
7. Los resultados de las pruebas diagnósticas y formativas aplicadas en las aulas por los docentes se presentan en función de niveles de desempeño. Estos resultados no sólo asisten al docente en identificar a estudiantes que se han quedado atrás y requieren atención, también identifican a aquellos en niveles avanzados que pueden servir como alumnos tutores.

APÉNDICE 1

Las siguientes cuatro gráficas sirven como ejemplo de la información que se reporta a nivel nacional en cada departamento en función de los resultados de las pruebas fin de grado aplicadas en 2010.

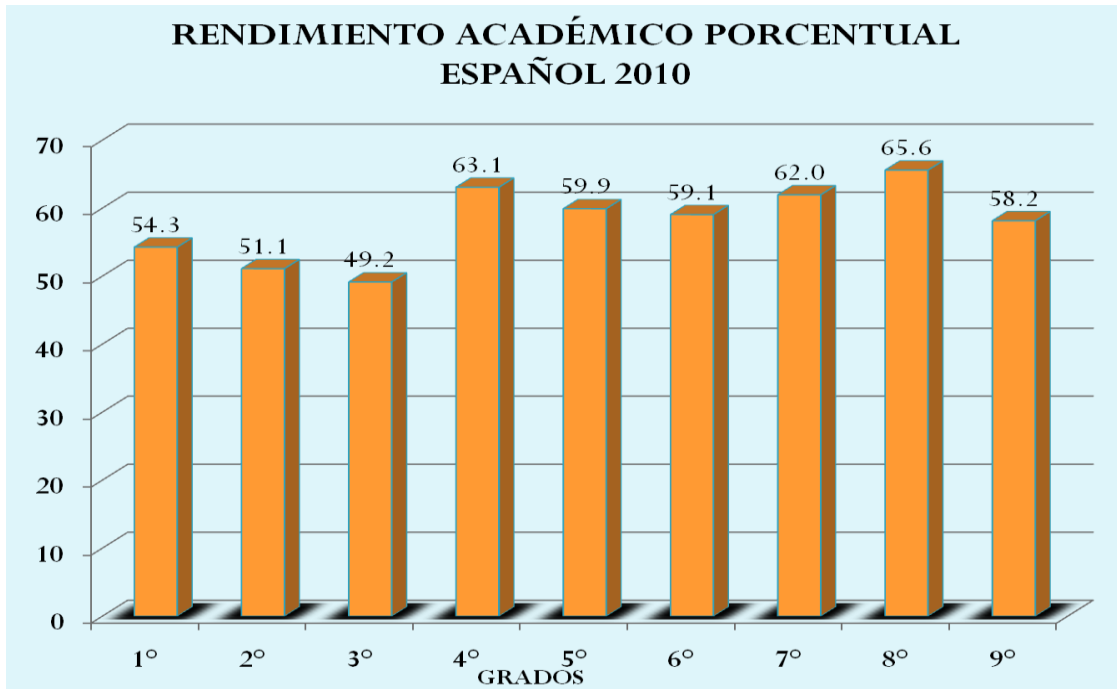
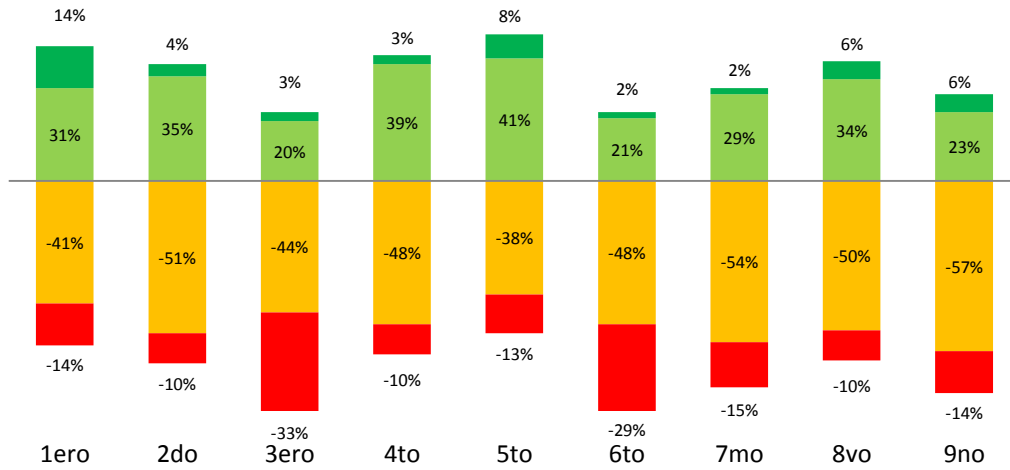


Figura A-1: Puntuaciones promedio para Español grados 1º-9º, porcentaje correcto.

Porcentaje de estudiantes por Nivel de Desempeño Español 2010



Insatisfactorio- Debe Mejorar – Satisfactorio- Avanzado

Figura A-2: Estudiantes por nivel de desempeño, Español grados 1º-9º.

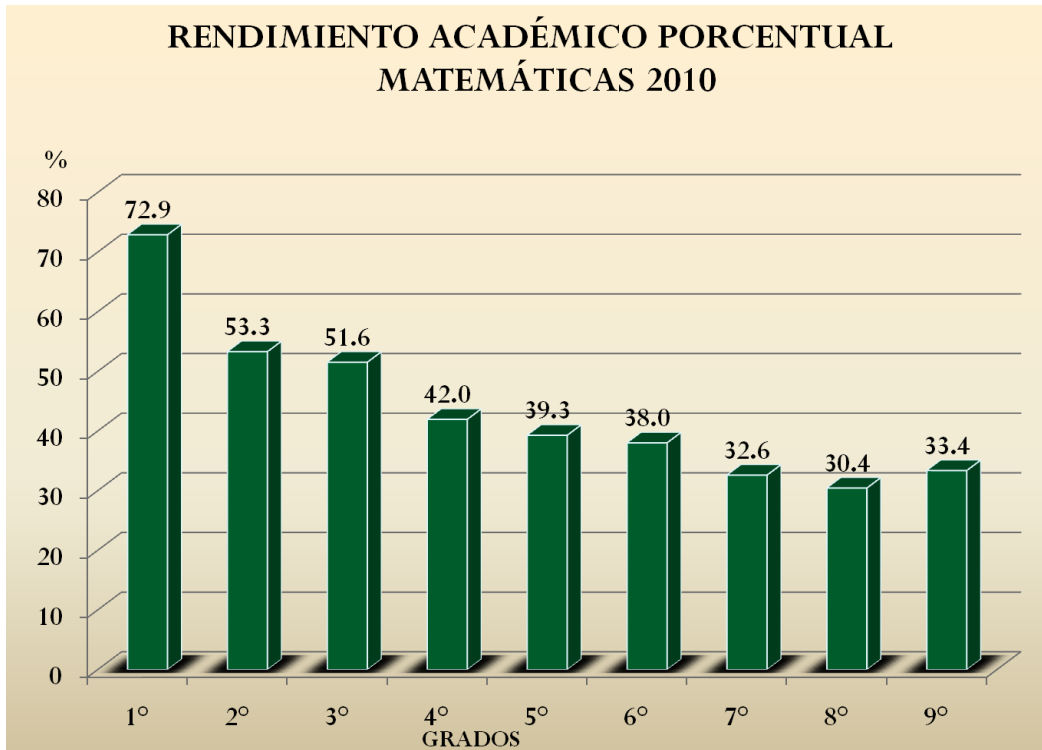
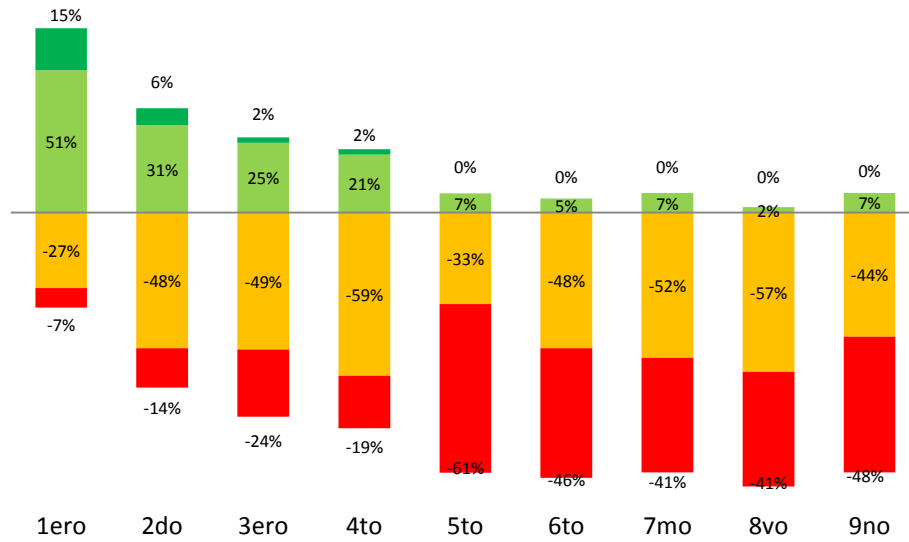


Figura A-3: Puntuaciones promedio para Matemáticas grados 1º-9º, porcentaje correcto.

Porcentaje de estudiantes por Nivel de Desempeño Matemáticas
2010



Insatisfactorio- Debe Mejorar – Satisfactorio- Avanzado

Figura A-4: Porcentaje de estudiantes por nivel de desempeño, Matemáticas grados 1º-9º.

APÉNDICE 2

MATERIALES EDUCATIVOS ALINEADOS AL DCNB



APÉNDICE 3

NOTAS:

Los informes de los resultados de las evaluaciones educativas externas fin de grado de los años 2007, 2008, y 2010 pueden encontrarse a través del sitio web de la Secretaría de Educación hondureña: <http://av.dcnbhonduras.org>

Los estándares, las programaciones, y las pruebas diagnósticas y formativas mensuales pueden encontrarse en el mismo sitio web. El informe de los resultados de la evaluación educativa externa de octubre/noviembre 2010 se presentó en el 2011, y presenta resultados hasta por cada estándar educativo, por área curricular, por grado, y por departamento, ofreciendo información valiosa para la planificación local de capacitación docente en base a necesidades reales.

La próxima evaluación externa se ha programado para octubre 2012. Los resultados de la evaluación educativa externa 2012 serán comparables con los resultados 2010, por el proceso de equiparación ya mencionado en el texto.

REFERENCIAS

- Alas, M., Hernández, R., Moncada, G. (2008). El CNB en el Aula: Disponibilidad, Uso, Valoración, e Impacto de los Materiales Educativos Alineados con el DCNB, 2006-2008. Informe Técnico, Proyecto MIDEH, Tegucigalpa.
- Car, J. F., & Harris, D.E. (2001). Succeeding with Standards: Linking Curriculum, Assessment, and Action Planning. Alexandria, VA: Association for Supervision and Curriculum Development.
- Cizek, G. J., Bunch, M. B., & Koons, H. (2004). Setting Performance Standards: Contemporary Methods. Educational Measurement: Issues and Practice.
- Equip 1 (USAID), (2012). Using standards and assessments to improve student learning outcomes in developing countries. Washington, D.C.: U.S. Agency for International Development and the American Institutes for Research.
- Glatthorn, A. (1999). Performance Standards Authentic Learning. Larchmont, NY: Eye on Education, Inc.

- Hamilton, L.S., Stecher, B.M., & Yuan, K. (2008). Standards-based reform in the United States: History, research, and future directions. Los Angeles, CA: RAND Corporation.
- Linn, R.L. (2001). The design and evaluation of educational assessment and accountability systems (CSE Technical Report 539). Los Angeles, CA: National Center for Research on Evaluation, Standards, and Student Testing.
- McMeekin, R. (2006). Hacia una comprensión de la Accountability Educativa y como puede aplicarse en América Latina. Santiago, Chile: Editorial San Merino.
- O'Shea, M. (2005). From Standards to Success. Alexandria, VA: Association for Supervision and Curriculum Development.
- Popham, W. J. (2006). Defining and Enhancing Formative Assessment. Assessment for Learning.
- PREAL. (2005). Informe de Progreso Educativo. Honduras 2005.
- PREAL. (2011). Educación: Un Desafío Impostergable. Honduras 2011.
- Secretaría de Educación de Honduras, web site: <http://av.dcnbhonduras.org>
- Zieky, M., & Perie, M. (2004). A Primer on Setting Cut Scores on Tests of Educational Achievement.