

Coordinación de



Estudios de Posgrado



# EL SENTIDO DIDÁCTICO DE LAS PRUEBAS A GRAN ESCALA: EL CASO DE ENLACE

TESTS

para obtener el grado de  
Doctora en Pedagogía

**Presenta**  
**Rosa Aurora Padilla Magaña**

Tutor:  
Dr. Ángel Díaz Barriga

Comité tutorial:  
Dra. Concepción Barrón Tirado  
Dra. Frida Díaz Barriga Arceo

Jurado:  
Dra. Martha Corenstein  
Dr. Marco A. Rigo Lemini.



Cd. Universitaria, D. F., a junio del 2010



Universidad Nacional  
Autónoma de México



**UNAM – Dirección General de Bibliotecas**  
**Tesis Digitales**  
**Restricciones de uso**

**DERECHOS RESERVADOS ©**  
**PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL**

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

**Un examen como instrumento de evaluación, sólo cobrará sentido, cuando su contenido esté vinculado con procesos de aprendizaje, que generen situaciones de enseñanza (RAPM)**



**Hagamos del examen más que un instrumento de evaluación, un instrumento de aprendizaje (RAPM)**







### **Algo que decir y mucho que agradecer....**

Por eso y más... en estas primeras líneas de una obra que representa, un esfuerzo, una meta, un aprendizaje, pero sobre todo, parte importante de un proyecto personal, no puedo dejar de mencionar, recordar y agradecer, a quienes han compartido esta inigualable experiencia. Este trabajo es de ustedes y para ustedes.

No se puede pensar en un proyecto de vida, si no se cuenta con ésta. *Lupita y Aarón* no sólo les debo la vida, sino mucho más... Gracias por enseñarme a saber vivir con plenitud y alegría, aceptando, valorando y disfrutando todo lo que la vida nos ofrece. Para aprender a trabajar incansablemente por un ideal. Pocas personas han sido tan aferradas a la vida como ustedes. Aunque partieron físicamente, su infinito amor, enseñanzas y vivencias compartidas, las tendré presentes en cada uno de mis pasos. Son los tesoros que me permitirán seguir adelante con esta difícil, pero invaluable tarea que me han dejado: continuar con mi proyecto de vida, viviendo intensamente hasta el final. Son y serán siempre mi ejemplo de vida, lucha y entrega.

Hay hermanos de sangre y hermanos de corazón. *Bel y Pepe*. Son los hermanos que la naturaleza me negó, pero que la vida me regaló. Saben que forman parte fundamental de mi existencia. Gracias mil, porque no importando el tiempo, la circunstancia o la distancia, siempre han estado ahí, para respaldar y apoyar. Su presencia en cada momento de la vida, ha sido refugio e inspiración. Sin compañeros como ustedes, alentando siempre en las adversidades y celebrando las victorias, las batallas de la vida, serían imposibles de salvaguardar.... Ha sido maravilloso recorrer juntos el camino.

Creer, ver de lo que somos capaces y atrevernos a volar, no es algo fácil de lograr. Gracias *Ángel*, por permitirme aprender y formarme bajo tu tutela. Como todo buen maestro, supiste sembrar inquietudes, abrirme un espacio de formación, colocarme frente a situaciones y retos que me permitieron descubrir posibilidades. Y finalmente, me ayudaste a volar. Sabes que eres parte importante de lo que hasta ahora he logrado. Tu presencia en mi formación académica y en momentos fundamentales de mi vida, han dejado una huella difícil de borrar.

En lo avatares de la vida, necesitamos de anclas para sostenernos y seguir adelante. *Canchita y Frida*, su constante apoyo, impulso y acompañamiento, no sólo en lo académico, sino en lo personal, ha sido fundamental para poder llegar al final del proceso. Su confianza y amistad, los pilares más importantes que me permitieron sostenerme y continuar...Gracias maestras y amigas por enseñarme a ver siempre el horizonte.

Muchos son los regalos de la vida pero una amistad sincera, es uno de las mayores riquezas que el ser humano pueda atesorar. Como olvidar todas las vivencias y experiencias compartidas con invaluable amigos como ustedes. Siempre dispuestos a apoyar, escuchar, impulsar y animar... Gracias, *Claudia, Coni, Caty, Arlen, Laura, Ido y Jorge Corlos*, por su inigualable forma de vivir la amistad. Porque siempre han estado ahí, por ser fieles compañeros de batalla.

Para crecer y realizarnos, siempre es fundamental el respaldo y apoyo de una familia. Gracias a todos los de la familia *IISUE-UNAM*, con quien he podido convivir durante mis años de formación, por acogerme, hacerme parte de ustedes y permitirme aprender. Por compartir conmigo no sólo un lugar o una vida académica, sino mucho más... Gracias por hacerse presentes, acompañar, e impulsar siempre....

Todo trabajo siempre es inacabable y perfectible. Gracias a la *Dra. Martha Corenstein* y al *Dr. Marco Rigo Lemini* por su disposición y ayuda para enriquecer este trabajo. Pero sobre todo por su confianza y apoyo permanente. Por todo lo que de ustedes he aprendido.

Contar con un espacio donde uno pueda colaborar y crear es parte importante en todo desarrollo personal y profesional. Gracias *Dr. Arnal, Jorge y Gaby*, por su confianza, apoyo y amistad. No se necesita mucho tiempo para reconocer y apreciar la gran calidez humana que los caracteriza. Es un orgullo y satisfacción formar parte de su equipo de trabajo.

A la familia *Padilla y Cabañas Saulés*, gracias por entender, respetar y apoyar. Por esos pequeños grandes detalles que hacen la diferencia. Por hacerme saber que cuento con ustedes.

*Para todos aquellos  
que han hecho de la educación  
una opción de vida.*

# ÍNDICE

	Pág.
<b>INTRODUCCIÓN</b>	
a. El problema del que partimos	I
b. La pregunta de investigación	VII
c. Propósitos de la investigación	VIII
d. Objeto de estudio	VIII
e. Construcción metodológica	VIII
<b>CAPÍTULO 1:</b>	
<b>SENTIDO Y FUNCIÓN DE LOS EXÁMENES A GRAN ESCALA. UN ACERCAMIENTO A SU TRANSFORMACIÓN Y ESTADO DEL ARTE</b>	1
<b>1.1 Gestación, desarrollo y transformación histórica del sentido del examen: exámenes escolares y a gran escala</b>	1
<b>1.2 El estudio de los exámenes a gran escala. Un acercamiento al estado de la cuestión</b>	12
1.2.1 Perspectivas teórico-metodológicas	12
1.2.2 Investigaciones recientes. Tendencias y aplicaciones	24
1.2.3 Modelos de análisis para los exámenes a gran escala. Un acercamiento a lo que se ha hecho	29
1.2.4 Producción e institucionalización de los exámenes a gran escala en el escenario nacional (1995-2007)	31
<b>1.3 La aplicación de exámenes a gran escala como parte de una política educativa:</b>	35
<b>1.4 Impacto sociopolítico y transformación del sentido de los exámenes a gran escala</b>	38
<b>CAPÍTULO 2:</b>	
<b>LOS EXÁMENES A GRAN ESCALA DESDE UNA PERSPECTIVA DIDÁCTICA</b>	41
<b>2.1 Una mirada didáctica para los exámenes a gran escala</b>	41
2.1.1 Reacciones y debates en el uso de los exámenes a gran escala como instrumento de evaluación	42
2.1.2 El examen como instrumento de evaluación desde un enfoque didáctico	47
2.1.3 Alcances y limitaciones de las pruebas a gran escala como instrumento de evaluación	49
2.1.4 Estudios y experiencias sobre un manejo didáctico de los exámenes a gran escala	52
2.1.5 Implicaciones en la construcción de una perspectiva didáctica para la evaluación del aprendizaje y los exámenes a gran escala	58

<b>2.2 Propuesta de un modelo didáctico de análisis de reactivos para los exámenes a gran escala</b>	63
2.2.1 Consideraciones básicas sobre la evaluación del aprovechamiento escolar.	63
2.2.2 Elementos para una fundamentación conceptual del modelo que se propone	64
2.2.3 Descripción conceptual de un modelo didáctico y sus componentes, para el análisis de reactivos de pruebas a gran escala	66
2.2.4 Variables de construcción del reactivo que afecta su resolución	78
<b>CAPÍTULO 3:</b>	83
<b>LA EVALUACIÓN DE LA COMPRESIÓN LECTORA EN LA PRUEBA ENLACE DESDE UN ANÁLISIS DIDÁCTICO</b>	
<b>3.1 Hacia una caracterización de la prueba ENLACE:</b>	84
3.1.1 Antecedentes y origen	85
3.1.2 Elementos que la caracterizan	83
3.1.3 Mecanismos de aplicación	86
3.1.4 Cobertura	88
3.1.5 Estrategias de operación	90
3.1.6 Proceso de elaboración y construcción de la prueba.	90
3.1.7 Elaboración y selección de reactivos	95
3.1.8 Intenciones explícitas, cuestionamientos que surgen y concepciones implícitas	102
<b>3.2 Un análisis didáctico de los reactivos de comprensión lectora de la prueba ENLACE 2008, para 6ª grado</b>	107
3.2.1 Generalidades de la prueba ENLACE 2008 para 6ª grado	107
3.2.2 Análisis didáctico de algunos de los reactivos de comprensión lectora en prueba ENLACE 2008	109
<b>3.3 Conceptualización y caracterización de la evaluación de la comprensión lectora en la prueba ENLACE</b>	131
<b>3.4 Alcances, limitaciones y sentido, de la estructura didáctica de los reactivos para evaluar la comprensión lectora en la prueba ENLACE 2008</b>	138
<b>CONCLUSIONES</b>	142
<b>REFERENCIAS</b>	151



## INTRODUCCIÓN

---

Evaluación y reforma, son dos conceptos, dos acciones y dos situaciones emanadas de los proyectos de modernización que sin duda alguna están marcando el rumbo de la educación básica en América Latina

La implementación de sistemas nacionales de evaluación, es parte de una tendencia mundial, que en la región latinoamericana, se ha convertido en el sello de la década de los noventa. Constituye un elemento central y una práctica común de los sistemas educativos que cada vez ha venido cobrando mayor fuerza a nivel nacional, regional y mundial, por las intenciones a las cuales responde, las funciones que se le atañen y las implicaciones de orden político, social, pedagógico y económico que éstas conllevan

### **a. El problema del que partimos**

De acuerdo con estudios realizados (véase Ravela, et al 2008), una de las mayores aspiraciones en la implementación de sistemas nacionales de evaluación en la región latinoamericana es la de mejorar tanto la adopción de políticas educativas, como la gestión de los sistemas educativos y el aprendizaje de los estudiantes.

Sin embargo, valdría la pena analizar en qué medida las evaluaciones a gran escala, están respondiendo a las necesidades y finalidades con las cuales fueron creadas y qué tipo de repercusiones previstas y no previstas conllevan, tanto en la sociedad, como en el sistema educativo y en sus actores

En materia de política educativa, la apuesta a la mejora de la calidad educativa, es el discurso en el que se ha venido instalando la construcción de toda esta serie de sistemas nacionales de evaluación en México y Latinoamérica

Sin embargo, consideramos que el auge de las evaluaciones a gran escala para medir el logro, determinar el grado de eficiencia de los sistemas educativos y mejorar su calidad, obedece más bien a la importancia otorgada por la sociedad actual, al uso de la evaluación para demostrar en qué medida los sistemas educativos están cumpliendo con el encomendado de que los estudiantes aprendan lo establecido en los diferentes niveles. Es decir, tiene una función más judicial y de verificación que de formación y mejora. Los esfuerzos realizados, que el conjunto de los países de la región han venido desarrollando para instalar sistemas nacionales de evaluación educativa, aún resultan insuficientes para obtener un sistema de evaluación que arroje la información necesaria para orientar el desarrollo de políticas educativas para el mejoramiento de la calidad de la educación.

La acción de definir e implementar mecanismos de evaluación externa en aras de la tan ansiada calidad de la educación, en un contexto como el latinoamericano, caracterizado por una serie de desigualdades no sólo económicas, sino sociales y educativas, conlleva tras de sí, una serie de situaciones y problemas comunes de diversa índole.

El empleo del sólo término "calidad de la educación", del que tanto se habla, pero cuyo sentido y significado es difícil de definir, representa en sí mismo un problema. Porque no existe una única definición. Tratar de conformar un sistema homogéneo con base en un estándar de "calidad de la educación", en países con situaciones culturales tan diversas, conlleva una dificultad adicional, de la que se derivan una serie de circunstancias sociales y educativas que vale la pena analizar.

Si bien se pensaba que por el sólo hecho de contar con un sistema de evaluación de la calidad, se iba a mejorar la educación, existen evidencias y estudios que demuestran lo contrario. La aspiración de una renovación educativa, que considere las grandes desigualdades sociales que afectan el desarrollo de nuestros países y dé respuesta a las necesidades culturales y retos educativos que nos plantea un mundo en constante cambio, se ha visto disminuida.

Si consideramos lo que la Conferencia Regional de Educación de la UNESCO para América Latina y el Caribe (OREALC/UNESCO Santiago) plantean en torno al significado de la calidad educativa, se hace necesaria una visión más amplia de este concepto, que más que una política educativa considere las dimensiones de orden social que éste conlleva. La calidad educativa, es una cuestión de equidad, relevancia y pertinencia, así como de tratamiento, uso y destinatarios de la información obtenida en las evaluaciones a gran escala.

Hablar de equidad, es tratar que todos los estudiantes tengan las mismas oportunidades educativas, ajustadas a las necesidades de cada uno de los destinatarios. Una educación de calidad requeriría entonces, atender a las desigualdades existentes. La homogenización que la evaluación estandarizada promueve, difícilmente considera la heterogeneidad y diversidad educativa y social existente. No se puede medir con la misma vara a quienes presentan situaciones y características socioeducativas y culturales diferentes. "Frente a las desigualdades sociales, la evaluación estandarizada puede reforzar situaciones de discriminación social" (LLECE, 2008, p.20)

En cuanto a la relevancia, hay que considerar que lo importante en la educación, está en relación al tipo de aprendizaje que promueve y a la medida en que éste responde a las necesidades sociales y al desarrollo personal. Lo cual tiene que ver mucho con la pertinencia, es decir con las diferencias existentes en el aprendizaje, que están en función de las características y necesidades de cada persona y el entorno sociocultural que las rodea. La evaluación de la pertinencia de la educación, debe considerar el tipo de aprendizaje evaluado y las posibilidades que éste promueve. Si consideramos que la mayoría de los exámenes a gran escala, son pruebas de lápiz y papel estandarizadas, con contenidos de tipo declarativo, que difícilmente reconocen las diferencias en el aprendizaje y las necesidades en torno a éste, la relevancia de la educación no puede ser evaluada mediante exámenes a gran escala, porque representan una mirada muy estrecha del conocimiento y del aprendizaje. En una región, caracterizada por la variedad de grupos étnicos con que cuenta, el logro de la calidad educativa y su valoración, implicaría implementar estrategias y acciones educativas y de evaluación, que consideren y retomen la diversidad existente en el conocimiento y situación de los estudiantes.

Desde esta perspectiva, la pertinencia de la educación y su evaluación, demandaría más que la exclusión, la inclusión de todas las poblaciones y sujetos. Sin embargo un sistema de exámenes sustentado en una lógica que busca la diferenciación del que sabe del que no sabe, con base en un estándar determinado, difícilmente podrá lograr la pertinencia y calidad de su sistema educativo. Las situaciones presentadas, dan cuenta de que los estándares, más que la calidad y equidad apoyan la desigualdad. Frente a los desafíos socioculturales de una región tan diversa, la evaluación de la calidad educativa, no puede reducirse a cuestionarios de mediciones masivas de tipo estandarizado.

Todo esto, lleva a pensar en la necesidad de desarrollar un sistema de evaluación que más que la clasificación y eliminación de quien no alcance los estándares establecidos, atienda las diferentes necesidades sociales, educativas y de aprendizaje de cada sociedad y contexto, ayude a los docentes

a identificar problemas en la enseñanza y el aprendizaje y promueva el desarrollo de aprendizajes significativos en los estudiantes

La evaluación para el mejoramiento de la calidad de la educación, más que un sistema de exclusión y diferenciación social, deberá considerar la necesidad de un conocimiento sobre los logros y dificultades de un proceso educativo y de aprendizaje. Esto requeriría de una evaluación centrada en el aprendizaje, más que en la verificación de un tipo de conocimiento específico. Sólo así se podría lograr que las evaluaciones acompañaran los procesos de generación de conocimiento y proporcionaran información, no únicamente de tipo cuantitativo, sino más bien cualitativo, a todos los involucrados de un sistema y en especial a los actores principales del acto educativo: los docentes.

Si consideramos que una de las aspiraciones en la implementación de sistemas nacionales de evaluación es la mejora no sólo de la calidad, sino de la gestión de los propios sistemas y del aprendizaje, esta intención difícilmente podrá verse alcanzada, cuando lo que se informa son sólo puntajes y no lo que éstos significan.

Todo sistema de evaluación parte del supuesto de que una vez conocidos los resultados alcanzados, éstos serán utilizados por todos los usuarios y se generarán de manera automática los cambios necesarios, tanto en la gestión como en la práctica. Sin embargo, esta intención raramente puede verse lograda, cuando no se considera a los destinatarios de esta información, el tipo de información que se proporciona, el imaginario que ésta produce en la sociedad y el uso que se le da a la misma. Éste consiste generalmente sólo en difundir los resultados y premiar o castigar con base en éstos, a los protagonistas.

La mayoría de las veces, los resultados se utilizan sólo para clasificar, juzgar y controlar. La evaluación masiva del rendimiento escolar, no siempre ha estado acompañada de una definición consensuada de lo que implica aprender, del tipo de aprendizaje que se promueve, así como tampoco de una claridad en la finalidad de estas acciones para orientar el uso y aprovechamiento de la información generada y con base en esto promover acciones educativas que más que premiar o castigar a los mejores o peores, ayuden a una real transformación de las prácticas docentes del aula.

Desde una mirada pedagógica, más que ayudar a mejorar los sistemas de enseñanza o transformar las prácticas educativas, la implementación de exámenes a gran escala, ha venido a deformar el trabajo del docente. Porque no hay un impacto en la transformación de la escuela, ni una vinculación con la

formación de docentes. Se ha visto que renovar la educación y optimizar la acción docente, no es cuestión de exámenes.

A pesar que en México contamos con más de cinco exámenes aplicados en diferentes momentos a los estudiantes,<sup>1</sup> lo único que con éstos ha ocurrido, es generar una polémica social y una gran desconfianza de la escuela a nivel nacional, a partir de la publicación de los resultados obtenidos.<sup>2</sup>

No se han generado aún las acciones encaminadas a mejorar lo que pasa en el interior de la escuela y en las aulas. Los exámenes no han ayudado a transformar la realidad. La función formativa de la evaluación, ha quedado atrás y más aún se ha venido deformando. Porque más que enriquecer la acción docente, la aplicación de exámenes a gran escala y lo que éstos generan, ha venido a empobrecer el trabajo de docentes y alumnos, transmitiendo una idea errónea de lo que significa aprender y enseñar, además de fomentar la corrupción de las prácticas escolares. Ahora lo más importante en la escuela, no es el desarrollo de las capacidades humanas y promoción de la cultura, sino pasar el examen. La vieja aspiración de la ilustración (Kant, 1979), de cultivar el espíritu y la mente a través de la educación ha quedado atrás.

Lo anterior resulta realmente preocupante, sobre todo si consideramos desde un plano económico, el costo financiero y no sólo educativo y social de las evaluaciones a gran escala como parte de un sistema de evaluación del aprendizaje. La información sobre los costos de las evaluaciones realizadas

---

<sup>1</sup> A nivel nacional, tenemos como mínimo tres exámenes estandarizados de tipo externo, que en diversos grados y por diferentes institutos se han venido aplicando a nuestros estudiantes. El Examen Nacional de Logro Académico en Centros Escolares, mejor conocido como ENLACE, prueba censal aplicada cada año a partir del 2006, por la Secretaría de Educación Pública, el Examen para la Calidad y el Logro Educativo (EXCALE), prueba muestral elaborado por el Instituto Nacional de Evaluación de la Educación (INEE) en el 2003 y el Examen Nacional de Ingreso EXANI-1, cuyos orígenes se pueden ubicar desde 1994, elaborado y aplicado al término de la educación secundaria por el Centro Nacional de Evaluación de la Educación Superior (CENEVAL). A nivel internacional son básicamente tres los estudios en los que México ha participado en los últimos 15 años: el Programa Internacional de Evaluación de Estudiantes (PISA), aplicado desde el año 2000 y cada tres años a alumnos de 15 años y los del Laboratorio Latinoamericano de Evaluación de la Calidad, que desde el año 1995 ha realizado dos estudios a nivel regional para evaluar a alumnos de tercero y cuarto grado de educación Básica, así como el Tercer Estudio Internacional sobre Matemáticas y Ciencias, (TIMSS), aplicado en nuestro país en 1995 y del cual México pidió que no se dieran a conocer los resultados. Sin embargo consideramos que a pesar de esta propagación de exámenes, la necesaria transformación y mejora educativa en las aulas, sigue siendo un asunto pendiente. En estudios anteriores (Padilla, 2006), presentamos un análisis comparativo del contenido de cada uno de estos exámenes.

<sup>2</sup> Sobre todo si tomamos en cuenta que por ejemplo, la prueba ENLACE, objeto de estudio de esta investigación, es un examen censal que desde el 2006 se ha venido aplicando consecutivamente cada año. Lo que implica que no hay tiempo suficiente para su revisión y sobre todo para el análisis de los resultados obtenidos y la creación e implementación de situaciones que ayuden en mejorarlo. Mientras pruebas internacionales como PISA (Programa Internacional de Evaluación de Estudiantes) tardan sólo 10 meses en el desarrollo de sus fundamentos, 11 meses en el desarrollo de los reactivos, 8 meses en la revisión nacional de los ítems (véase OECD, 2005). En México, para poder aplicar anualmente la prueba ENLACE, sólo se ocupan de 2 a 6 meses para su elaboración y menos de 6 meses para su corrección.

en cada país, es aún escasa. En América Latina, un estudio preliminar sobre el tema (Wolff, 2007), reporta la información proporcionada por funcionarios encargados de los programas de evaluación de cinco países de la región: Chile, Colombia, Honduras, Perú y Uruguay, sobre los costos de sus evaluaciones. Los datos que se obtuvieron fueron indicativos, más nunca definitivos. Dado que las condiciones de evaluación difieren mucho de un país a otro, los costos también varían. Además éstos, dependen del número de alumnos evaluados y del tipo de evaluación (nacional o internacional, muestral o censal) aplicada. Según datos arrojados por este estudio, los costos van desde \$2.50 dls por alumno en una evaluación de más de un millón de estudiantes en Colombia, hasta alrededor de \$100 dls por alumno en Perú y Colombia por una evaluación muestral de 5,000 alumnos en una prueba internacional (Wolff, 2007, p.8).<sup>3</sup>

Muchas son las implicaciones y consecuencias en cuanto a los usos que se le han atribuido a las evaluaciones a gran escala. Sin embargo la mayoría de las veces, éstas poco tienen que ver con lo educativo. Aunque un grupo de expertos en evaluación de la región<sup>4</sup> a través del Programa de Reforma Educativa en América Latina (PREAL), han publicado diversos estudios sobre la situación de las evaluaciones educativas a gran escala, en América Latina, el impacto que este tipo de evaluaciones pueda tener, sobre todo en los procesos del aula, sigue estando ausente.

Sirva lo anterior para apuntalar las acciones y consecuencias de orden político, social, educativo y económico de un sistema de evaluación aún no previstas y la imperiosa necesidad de transitar de un manejo meramente político-administrativo, a un asunto más pedagógico y didáctico, en el abordaje de las pruebas a gran escala, como instrumentos de evaluación.

Esto nos lleva a pensar en la necesidad de reflexionar y analizar el rumbo y sentido que la evaluación del aprendizaje, sobre todo en su modalidad a gran escala, está teniendo hoy y sobre el que debe adquirir.

El sentido de la evaluación educativa, de la enseñanza y del aprendizaje, así como de la tan ansiada mejora educativa está en riesgo. La evaluación del aprendizaje al materializarse en exámenes a gran

---

<sup>3</sup> En cuanto a México se refiere, de acuerdo con información proporcionada por Josefina Vázquez Mota, la entonces titular de la SEP durante el periodo de noviembre 2006 a abril del 2009, para el ciclo escolar 2008, la prueba ENLACE tuvo un costo superior a 220 millones de pesos (Poy, 2008).

<sup>4</sup> Entre estos podemos citar a Pedro Ravela de Uruguay, Patricia Arregui de Perú, Richard Wolfe de Canadá, Guillermo Ferrer y Margarita Poggi de Argentina y Lawrence Wolff de Estados Unidos.



escala, ha dejado de ser parte del proceso didáctico, para convertirse en un elemento más de política educativa y rendición de cuentas. En lugar de ser un recurso del docente para que los estudiantes demuestren lo aprendido y en caso de error, lo corrijan y generen un mejor aprendizaje, a partir de lo cual los mismos docentes, podrían revisar su acción didáctica, el examen y el peso que a éste se le ha dado, en relación a los resultados, se ha convertido hoy en día, en el fin de la acción de profesores y alumnos y en un instrumento contra el docente. Es decir, se ha alejado del aula.

De ahí nuestro interés por a través de este estudio, devolverle a la evaluación de aprendizaje, sobre todo en su modalidad a gran escala su enfoque formativo y sentido didáctico. Consideramos que para que la evaluación cumpla con su función formativa de mejorar la enseñanza y el aprendizaje, es necesario pasar de la evaluación como objetivo, o más aún como indicador, a la evaluación vinculada al proceso de aprendizaje, al contenido de evaluación y a las formas de enseñanza. En donde docentes y sociedad en general, más que los resultados de estas evaluaciones, conozcan y participen de la finalidad y contenido de los exámenes a gran escala que nuestros estudiantes están presentando. Lo que demanda un análisis de su estructura didáctica y de sus implicaciones en materia educativa.

#### **b. La pregunta de investigación**

Ante la construcción y administración de exámenes a gran escala, con una perspectiva meramente cuantitativa, caracterizada por una visión administrativa y con fines clasificatorios, que coloca al examen como instrumentos de medición, lejano a la comprensión de los procesos de aprendizaje y del sentido formativo que debe perseguir la evaluación del mismo, vemos la necesidad de proponer un sistema de análisis de reactivos para exámenes a gran escala que desde una mirada didáctica, permita recobrar el sentido formativo de la evaluación del aprendizaje y dar cuenta, tanto de los procesos de enseñanza-aprendizaje implícitos en su elaboración y resolución, así como de los alcances o limitaciones educativas.

Por la magnitud de su aplicación y por lo que esto conlleva, tomamos como referente la *Evaluación Nacional de Logro Académico en Centros Escolares*, mejor conocida como prueba ENLACE, en especial, la evaluación que a través de esta se realiza de un área básica como el español, y de la comprensión lectora. Así, la pregunta que da pie al siguiente estudio es

*¿Qué puede revelar un análisis de corte didáctico en los reactivos de español, de pruebas a gran escala como ENLACE?*

**c. Propósito de la investigación:**

El propósito de esta investigación es ofrecer un modelo de análisis de reactivos, que desde una perspectiva didáctica, permita conocer el contenido de evaluación de una prueba a gran escala en una determinada área de evaluación y lo que éste implica en materia educativa

A partir de un análisis de los elementos que intervienen en el proceso de aprendizaje del alumno y su vinculación con situaciones de enseñanza, se pretende darle un abordaje didáctico a lo que constituye una prueba y lo que a través de ésta se puede conocer, más allá de los resultados

**d. Objeto de estudio.**

De la problemática planteada y los propósitos establecidos, el objeto de estudio de esta investigación es

*Analizar desde una mirada didáctica, el contenido de un área de evaluación de una prueba a gran escala, como ENLACE.*

**e. Construcción metodológica**

La realización del presente trabajo siguió varias fases, las cuales sirvieron como base para la construcción de los capítulos que comprende

Una de las primeras etapas consistió en la revisión del estado del arte en el que se encuentran los exámenes a gran escala. Esto permitió identificar las corrientes de pensamiento desde las cuales se han estudiado a los exámenes a gran escala y acercarnos el estado que guarda la producción sobre exámenes a gran escala en México en un periodo de doce años (1995-2007). La información recabada

fue la base para la construcción del primer capítulo denominado "El sentido y función de los exámenes a gran escala. Un acercamiento a su función y estado del arte."

A partir de la revisión de la literatura existente para la construcción del estado del arte en el que se encuentran los exámenes a gran escala donde queda de manifiesto la ausencia de un abordaje didáctico en su tratamiento. A partir de las reacciones y debates que el examen ha venido generando como instrumento de evaluación, los alcances y limitaciones que éstos presentan y los estudios y experiencias que se han podido identificar para un manejo didáctico del examen, se propone la construcción de una perspectiva didáctica para los exámenes a gran escala y sus implicaciones. Esto lleva a proponer la construcción de un modelo para analizar los reactivos de exámenes a gran escala, desde un enfoque didáctico, donde se presenta una conceptualización y descripción de los elementos que lo componen, considerando además las variables en la construcción del ítem que afectan su resolución. Siendo así como se conforma el segundo capítulo del trabajo titulado "Los exámenes a gran escala desde una perspectiva didáctica."

Como referente empírico para la aplicación del modelo propuesto y ante la imposibilidad de analizar toda la prueba, compuesta por 178 reactivos, tomamos como referencia el 30% de los reactivos del área de español de la prueba ENLACE, 2008 para 6° grado. Los reactivos fueron seleccionados con base en el grado de dificultad que se les asignó estadísticamente (SEP, 2008). Esto con el objeto de conceptualizar y caracterizar el contenido de una de las áreas más comunes de evaluación en pruebas estandarizadas y ver en qué medida el grado de dificultad establecido, corresponde a la dificultad real presentada en cuanto a su contenido y estructura. El análisis realizado, constituyó el tercer capítulo de esta investigación, el cual permitió dar cuenta de las características que guarda la evaluación de la comprensión lectora en pruebas a gran escala como ENLACE y del sentido que ésta adquiere en relación a sus alcances y limitaciones encontradas.

Consideramos que lo más importante del trabajo realizado, es que ayudó a encontrar una alternativa para abordar el contenido y estructura de los medios con los que se pretende valorar el aprendizaje y el trabajo de los docentes. Lo que a su vez ayudó a reflexionar sobre la congruencia o incongruencia existente entre lo que se pretende evaluar a través de estas pruebas y el tipo de aprendizaje y finalidades educativas que la escuela de hoy está demandando.

La investigación que aquí se presenta, da cuenta también de mensajes implícitos que en relación a la evaluación, al aprendizaje y al sentido que éstos adquieren cuando se pretende evaluar o más bien medir el aprendizaje, a través de pruebas a gran escala como ENLACE

## Capítulo 1

### SENTIDO Y FUNCIÓN DE LOS EXÁMENES A GRAN ESCALA. UN ACERCAMIENTO A SU TRANSFORMACIÓN Y ESTADO DEL ARTE

*“Las pruebas con frecuencia no recogen información adecuada sobre el desarrollo intelectual y personal que están experimentando los estudiantes, no son más que juegos en los que se gana o se pierde” (Bain).*

El examen, se ha convertido en uno de los instrumentos más utilizados para la evaluación y/o medición del aprendizaje y aprovechamiento escolar. Desde su aparición, en el ámbito educativo, ha tenido diversas funciones y sufrido algunas modificaciones. Una de las más importantes es la transfiguración de su sentido y el manejo social, más que pedagógico que a éste se le ha dado. El examen a gran escala se ha convertido hoy, en un instrumento de política educativa para juzgar y controlar a los sistemas educativos, lo que lo ha desvinculado del aprendizaje.

La transformación del sentido del examen y el papel que éste ha jugado, se puede entender a partir de una revisión del surgimiento y evolución del examen en el ámbito educativo. Podemos distinguir dos etapas y modalidades: la de los exámenes escolares a manera de ensayo y la de los exámenes a gran escala, cuyo desarrollo ha pasado por diferentes momentos. Éstos han sido estudiados desde diferentes corrientes de pensamiento y perspectivas. Para su análisis se ha recurrido a diferentes modelos.

La producción generada en México sobre evaluaciones a gran escala, durante los últimos doce años permitirá conocer el estado que éstos guardan a nivel nacional, dar cuenta de su institucionalización y comprender su impacto social en el ámbito nacional.

#### **1.1 Gestación, desarrollo y transformación histórica del sentido del examen: exámenes escolares y exámenes a gran escala**

Históricamente el sentido y función del examen tanto escolar como a gran escala, se ha transformado y ha adquirido diferentes vías.

Desde su nacimiento en la China imperial hace más de tres mil años, como una forma de seleccionar funcionarios para puestos importantes del gobierno, se le han atribuido funciones políticas y sociales más que educativas.

Su aparición en el ámbito educativo, ha obedecido a diversos intereses y transcurrido por diferentes etapas, formas y usos. Una revisión de la finalidad con que los exámenes incursionan al ámbito educativo y de cómo ésta se ha ido modificando, permitirá abordar el problema de los exámenes, primero a nivel escolar y más tarde en su modalidad gran escala, en cuanto a la función política-social, más que educativa que éstos han adquirido.

Las primeras aplicaciones de exámenes en materia educativa, las podemos encontrar durante la Edad Media, cuando la Universidad Medieval inició con la presentación de exámenes como una manera de demostrar madurez académica y acceder el título de bachiller, la *licencia docendi* o el grado de doctor (Durkheim, s/f). Se trataba de exámenes orales realizados a manera de "Discusión pública" denominados *déterminance*, cuya finalidad era el planteamiento y defensa de tesis. Cabe señalar que en esta época los aspirantes sólo podían presentar examen cuando se estaba seguro de su preparación y éxito. El examen era considerado entonces, sinónimo de madurez académica.

Más tarde, la escuela Jesuita, en el siglo XVI, concedió gran importancia a los exámenes escritos, como una forma de promoción y en su *Método de estudios inferiores*, conocido, como *Ratio Studiorum* (Labrador, et al., 1986) se establecieron uno de los primeros reglamentos para la realización de exámenes tanto orales como escritos. Con lo cual, el examen comienza a ritualizarse y considerarse un instrumento de penalidad y competencia.

En el siglo XVII, Juan Amos Comenio, los introduce como parte del método de la enseñanza, con la finalidad de revisar y corregir lo aprendido.

Dice Comenio (1922,2000)

"Para examinar los trabajos al dictado, si los hay, ordenar a uno, dos, o más si es necesario, que lean lo que han escrito con voz clara y distinta, señalando expresamente las pausas, y entretanto todos los demás harán las correcciones mirando sus cuadernos. Para los ejercicios de traducción de una lengua a otra *todos examinarán* sus cuadernos y corregirán lo que hayan equivocado" (p. 100)



La desvinculación del examen con el método de enseñanza, ocurre durante el siglo XIX, cuando el examen aparece vinculado a todo un sistema de calificaciones, lo que cambia el sentido de la enseñanza y el aprendizaje, al centralizar las acciones en torno a un examen y matar el deseo de aprender en los estudiantes<sup>5</sup>

De ahí que Foucault los haya concebido como un instrumento a través del cual, las relaciones de saber se invierten por las relaciones de poder, dándole un uso normalizante. A esto se debe que en todos los dispositivos de disciplina, el examen se encuentre altamente ritualizado. Y sea utilizado como un medio para vigilar, sancionar y castigar (Foucault, 1977)

El Cuadro 1 muestra la evolución histórica que ha tenido el examen escolar y las distintas funciones que éste adopta en cada época

Cuadro 1

<b>El examen escolar. Desarrollo histórico</b>			
<b>Época</b>	<b>Siglo</b>	<b>Características y usos</b>	<b>Función</b>
Antigua		Origen en la dinastía China como medio de selección	Seleccionar
Medieval	XII	Aplicación de exámenes de corte escolástico en la universidad para determinar la idoneidad de los candidatos y poner a prueba los conocimientos adquiridos	Demostrar lo aprendido y obtener un grado académico
Escuela Jesuita	XVI	Inicio de exámenes escritos, y elaboración de los primeros reglamentos para la elaboración de exámenes	Ritualizar
Comeniana	XVII	Exámenes vinculados al método de enseñanza	Revisar lo aprendido
	S XIX	Vinculado a un sistema de calificaciones	Calificar
	S XX	Instrumento de poder y normalización	Corregir y mantener el orden

Fuente: elaboración propia

El Cuadro 1 permite observar cómo el examen desde su aparición e incursión al ámbito educativo, ha tenido distintas funciones, que poco o nada tiene que ver con el aprendizaje. Sólo Comenio los

<sup>5</sup> Giner de los Ríos (s/f) en su libro "0 educación o exámenes", revela la transfiguración pedagógica que ha sufrido el examen por su modus operandi

presenta vinculados al método de enseñanza como una herramienta del maestro para revisar lo aprendido. De ahí que desde su ritualización y normalización, haya recibido fuertes críticas y cuestionamientos sobre la finalidad y valor educativo de los mismos. Para Hernández Ruiz (1949) “el examen por sí mismo no tiene valor. Hay que ver qué se examina, cómo se examina, con qué grado de frecuencia, para qué y con qué planes valorativos” (p. 81).<sup>6</sup>

Sin embargo, con el auge de la ciencia y la tecnología y la fuerte influencia que ejerció el pensamiento positivista<sup>7</sup> en materia educativa, el problema de la transformación del sentido educativo de los exámenes se acentúa cuando a finales del siglo XIX y principios del XX, surgen los primeros exámenes a gran escala, dentro del movimiento pragmático estadounidense.

El nacimiento de los exámenes a gran escala obedece tres acontecimientos ocurridos en diferentes circunstancias:

En un primer momento, la idea de aplicar exámenes a un gran número de estudiantes nace como parte del proceso de transformación que sufre la escuela estadounidense durante su proceso de industrialización.<sup>8</sup> Rice, uno de los promotores principales de este movimiento, inicia la aplicación de pruebas a gran número de estudiantes en 1894. Se trataba de una prueba de ortografía —que no estaba aún estructurada desde la teoría del test— aplicada a 16 000 alumnos.<sup>9</sup> La idea era apreciar las diferencias existentes en el rendimiento y la eficacia de los métodos para su enseñanza y con ello sustentar una protesta para impulsar su programa reformista, esto con la finalidad de conocer la situación en materia de aprendizaje y su relación con la enseñanza.

<sup>6</sup> Hernández Ruiz (1972) en su libro *La Crisis de la educación contemporánea*, refiere además 12 ideas por las cuales se rechazaba la aplicación de exámenes.

<sup>7</sup> La incursión del pensamiento positivista en materia de elaboración y aplicación de exámenes, da cuenta de la fuerte influencia sociológica que presenta el uso de exámenes para medir lo educativo. Desde un enfoque positivo, la ciencia concebida como la representación formal del objeto dado, tiene que ser representada forzosamente a través de un código matemático para cuantificar sus resultados y establecer leyes o regularidades entre los mismos. Es decir, se trata de un proceso de acumulación de datos, sobre un objeto observado y en donde la cuantificación de un hecho depende de la precisión obtenida bajo un método cuantitativo. Desde esta postura, la expresión numérica de un fenómeno o una relación, es garantía de exactitud en la observación, de rigor en la formulación de la ley y de precisión en los límites de la predictibilidad. Principios altamente discutidos por Benedito (1986), quien a partir del análisis de los sustentos de esta orientación filosófica, caracterizados por un modelo acumulativo, basado en la observación de hechos para la formulación de leyes, con el fin de predecir, en aras del progreso, cuestiona y discute el problema de la medición psicológica.

<sup>8</sup> El avanzado desarrollo industrial en que había entrado EEUU, ocasionó que a partir de 1840, en la segunda era de la Revolución Industrial, la estrecha vinculación empresa-educación, desviara al sistema norteamericano del humanismo y el ímpetu de las teorías administrativas, incursionaran al ámbito educativo, a partir de lo cual el eticentismo pragmático se adopta como modelo educativo de este país.

<sup>9</sup> La aplicación de estas pruebas estuvo precedida por visitas a las escuelas, conversaciones con los maestros, asistencia a las reuniones de las juntas de educación y entrevistas a los padres, pues su meta era ofrecer a público una evaluación objetiva (Cremin, 1969).

El segundo momento, tuvo como antecedente la medición psicológica, ocurrida en Francia a principios de siglo XX, con la creación de los tests de Binet y Simon<sup>10</sup> para medir la inteligencia. Dos psicólogos, que fueron los pioneros en la idea de construir una escala para medir la inteligencia como una serie de problemas de dificultad graduada, cada uno de los cuales corresponde a la norma de un distinto nivel mental. Lo que ayudó a sentar las bases para la aplicación de escalas a cuestiones de rendimiento. Éste es el momento donde surge la construcción del test, aunque no de manera masiva, ni bajo los principios de la teoría del test, como instrumento objetivo para medir una cualidad, que en ese momento era la inteligencia y posteriormente aprendizaje.

En un tercer momento dentro del movimiento pragmático estadounidense, ubicamos a Edward Thorndike, autor al que se le reconoce como pionero en la medición del rendimiento escolar. Sus planteamientos con respecto al aprendizaje y a la educación, establecieron las bases de la medición educativa para la elaboración de test. Para Thorndike el aprendizaje consistía en un cambio de conducta producido por la asociación entre un estímulo y una respuesta y en donde de acuerdo con su "Ley del efecto" los resultados eran la base para este aprendizaje. La magnitud del acto educativo dependía de la medición de sus resultados y lo que interesaba en educación era medir los cambios producidos.<sup>11</sup> Ésta es sin duda la etapa en la que la idea de medición del rendimiento con una visión cientificista y eficientista generada por Thorndike incursiona en el ámbito del aprendizaje y la educación. El objetivo de Thorndike era la creación de una ciencia pedagógica sobre la cual pudiera fundarse toda la educación (Cremin, 1969). Una nueva idea de conocimiento centrado en el dato y en su comprobación empírica, donde lo "científico" es concebido como una representación de lo real y exacto, validado a través de códigos matemáticos, estaba imperando en la educación estadounidense y en la ratificación de la misma a través de exámenes estandarizados.

La transformación educativa ocurrida en Estados Unidos a inicios del siglo XX, inspirada en una filosofía pragmática donde un pensamiento sólo es comprensible en relación con la práctica (James,

---

<sup>10</sup> Cabe señalar que Binet al iniciar el desarrollo de un instrumento de medida "válido", "objetivo" y "práctico" para identificar las capacidades mentales de los que presentaban dificultades en el aprendizaje, fue uno de los primeros en trasladar los conceptos de medición a la psicología.

<sup>11</sup> Así, bajo un enfoque conductista del aprendizaje y con la idea de que el aprendizaje era un producto que se podía medir Thorndike y sus discípulos comenzaron durante los primeros 15 años del siglo XX a elaborar escalas para medir el rendimiento en aritmética, (1908), escritura (1910), deletreo (1913), dibujo (1913), lectura (1914) y aptitud lingüística (1916). Considerándose ésta, la etapa precursora de los tests, estandarizados.

1975)<sup>12</sup> transformó no sólo la idea de escuela, sino el sentido de muchas de sus prácticas como lo es la aplicación de exámenes<sup>13</sup>

El Cuadro 2 muestra los diferentes acontecimientos que dieron origen a la generación de exámenes a gran escala

Cuadro 2

**Gestión de los exámenes a gran escala**

Época	Lugar	Impulsores	Acontecimiento	Finalidad
Finales siglo XIX 1894	Francia	Influencia sociológica Pensamiento positivista	Auge ciencia y tecnología	Cuantificación de resultados Establecimiento de regularidades entre los mismos
	EEUU	Rice	Transformación de la escuela estadounidense	Impulsar una reforma educativa inspirada en la filosofía progresista estadounidense
Principios del siglo XX	Francia	Binet y Simon	Medición psicológica a través de test	Construcción de escalas para medir la inteligencia
	EEUU	Thorndike	Medición del rendimiento escolar	Contar con una base científica para medir eficiencia escolar como parte del movimiento pragmático estadounidense

De acuerdo con lo presentado en el Cuadro 2, tenemos que en el marco de un reformismo educacional estadounidense, la aparición de exámenes a gran escala, estuvo vinculada a la idea de medir "científicamente" los resultados del aprendizaje, con el propósito de buscar la eficiencia escolar. Más que conocer la situación de aprendizaje, lo que interesaba era contar con una escala validada estadísticamente para medir, comprobar y comparar el rendimiento de los estudiantes

<sup>12</sup> Para James (1975), un significado que no sea práctico carece de sentido, sólo se puede adquirir una comprensión perfecta del pensamiento, al preguntarnos los efectos prácticos que puede tener sobre el objeto. Sin embargo el pragmatismo es una filosofía que conjunta al racionalismo, con el empirismo, que es religiosa como el racionalismo, pero al mismo tiempo como el empirismo, conserva el contacto con los hechos

<sup>13</sup> La reforma educacional estadounidense centrada en el progresismo, considera a ésta una oportunidad para promover su causa y según palabras de Dewey pionero en la educación progresiva, orientar la readaptación educacional, aunque ésta tuvo distintas entonaciones. Del progresismo social se pasó a la pedagogía centrada en el niño y después a un progresismo político que eclipsó en cierta forma a la educación progresiva. (Cremin, 1969)

La necesidad de medir y la adopción de un lenguaje fiscalista en materia de medición psicológica, incursiona al ámbito educativo sobre todo en lo que respecta a la elaboración de exámenes. Ayer (1965), afirma que "toda proposición de psicología puede formularse en lenguaje fiscalista, porque todas las proposiciones de psicología describen acontecimientos físicos" (p. 171). La física a través de su método experimental y desde una postura empirista, se convierte en el modelo para hacer ciencia. Así, la adopción del lenguaje fiscalista, se ha convertido en el lenguaje universal de la evaluación hoy. A través de este lenguaje se puede convertir a cualquier sistema en una ciencia física, "una ciencia física, cuya misión consiste en describir sistemáticamente la conducta física de los seres vivientes, en especial la de los humanos y en formular las leyes bajo las cuales puede subsistir esa conducta" (Ayer, 1965, p. 195). Este modelo ha sido rápidamente adoptado por las ciencias sociales, en especial por la educación y la evaluación.

Así, la física, su lenguaje y pensamiento, se convierten en una influencia importante para la medición educativa. En donde con el afán de convertir también a la educación en una ciencia exacta, se vio la posibilidad de medir el conocimiento a través de test. Bajo la creencia de que los conceptos de las ciencias, en especial de la física, eran los más óptimos para realizar mediciones educativas y tratar de expresar en unidades tangibles, los elementos intangibles (Tyler, 1967) se piensa que sólo es científico, aquello que puede ser observable, cuantificable y por ende medible.

La incorporación de la Teoría del test al campo de la evaluación, trata de científizar el acto evaluativo y suplantar a este concepto por el de medición. La aparición de los tests, en el campo de la psicología y la evaluación, trató de darle objetividad, validez y científicidad al acto educativo. Sin embargo cabe señalar que son precisamente las cualidades que caracterizan a un test, como pruebas validadas estadísticamente y elaboradas bajo una situación controlada, las que cuestionan la científicidad que se le pretende dar a la evaluación del aprendizaje a través de test.

De acuerdo con Benedito (1986), es importante considerar que "una disciplina no adquiere científicidad por el uso de un método general propio para todas las ciencias" (p. 179). Sino que el método con el que trabaja una ciencia depende de su objeto de estudio. Si consideramos que el objeto de la evaluación-medición de las pruebas a gran escala es el aprendizaje, éste, dado los fundamentos de una prueba, es considerado como una situación, homogénea, controlada, objetiva y siempre observable, que

presenta patrones o unidades iguales para su medición y comparación. Aquí surge la pregunta ¿realmente el aprendizaje es un objeto susceptible de ser medido a través de escalas?

La meta final de Thorndike de poner en marcha una ciencia pedagógica para medir la efectividad del hecho educativo tuvo y sigue teniendo gran auge a través de los sistemas de exámenes a gran escala. Su desarrollo ha tenido fuerte influencia, primero de la sociología a través del pensamiento positivo, luego de la psicología, con la adopción del lenguaje fisicalista y la incorporación de la teoría del test, para la generación de exámenes a gran escala, cuya incursión al campo educativo la podemos ubicar en diferentes etapas:

**a) La etapa precursora (1900-1915)**

Ésta ocurre cuando Thorndike y sus discípulos comenzaron a elaborar escalas para medir el rendimiento en aritmética (1908), escritura (1910), deletreo (1913), dibujo (1913), lectura (1914) y aptitud lingüística (1916).

**b) La etapa de auge (1915-1930)**

El periodo comprendido de 1915 a 1930 se distingue como la etapa de auge, por la diversidad de elaboración y aplicación de baterías de test que se elaboraron para todas y cada una de las destrezas escolares. Es la etapa de los indicios de la normalización de los tests, es decir del establecimiento de normas de aprovechamiento para una edad o grado con el fin de comparar el rendimiento de cualquier individuo.

El movimiento del test en su versión originaria, ocurrida durante los primeros treinta años del siglo XX, Tyler (1967), fue el periodo en el que los elaboradores de test centraron su atención en las mejores formas para la medición "objetiva" de los diferentes aspectos, momento que se caracterizaba por la atención puesta, tanto en el porcentaje de alumnos que contestaban acertadamente un test y la construcción de formas comparables en el mismo test, como a la importancia de que el contenido de los tests, tuviera una significación funcional. Sin embargo esto no se logró del todo, pues el contenido se manejaba conforme a los libros de texto que más se utilizaban, así como al contenido de los diversos planes de estudio, o bien, la selección se hacía conforme a las cuestiones que con mayor frecuencia se incluían en los exámenes estatales.



Esta etapa es reconocida como la del desarrollo de los test, por la proliferación que se dio a la aplicación de baterías de test. Es la época denominada por algunos autores (Escudero, 2003) como la del "Testing" porque la medición realizada a través de test de rendimiento, eran sinónimo de evaluación, situación que prevalece hasta la fecha.

**c) La etapa de crítica y cuestionamiento (1940-1950)**

Después de un período de auge, en materia de test estandarizados, denominado como el tayloriano, en donde lo que se buscaba era la eficiencia en la medición educativa, viene una etapa de decadencia. Es el momento de crítica y cuestionamiento al movimiento de los tests, en donde se destaca la necesidad de cuidar el uso que se les daba, pues se había visto que en virtud de que no alcanzaban los resultados deseados en el programa, no podían ser considerados como indicador de eficiencia de un sistema educativo. Así mismo, había que cuidar las conclusiones obtenidas de los datos arrojados por los tests y pensar el peligro que se generaba al pensar en el movimiento de los tests como una panacea para los problemas educativos. Se les reprochaba a los tests su centralización en objetivos estáticos, lo que acentuaba el status quo de la educación. Además se vio que en vez de ayudar a los maestros a mejorar su tarea, éstos provocaban que los esfuerzos de la escuela, se concentraran en los aspectos medidos por los tests. Lo que los ubicaba como un medio para imponer uniformidad al sistema (Tyler, 1967).

**d) La etapa de internacionalización (1950-2000)**

La más reciente era de la medición educativa a través de exámenes a gran escala, la podemos denominar como la era de la internacionalización de las pruebas masivas. Ésta comienza cuando a mediados del siglo XX, la *Internacional Association for the Evaluation of Educational Achievement* (Asociación Internacional para la Evaluación del Logro Educativo) (IEA)<sup>14</sup> con sede en Ámsterdam, conforma un consorcio cooperativo independiente de institutos y agencias de investigación en más de 45 países, encargada de realizar mediciones y estudios

---

<sup>14</sup> Su sitio en internet es: <http://www.iea.nl/>

de contexto orientados a conocer el nivel y condiciones en que se da el aprendizaje de las matemáticas y las ciencias, en los estudiantes del nivel básico y medio superior<sup>15</sup>

A este primer intento le siguieron otros trabajos que pretendían un análisis comparativo a nivel internacional, como los del *Educational Testing Service* (ETS)<sup>16</sup> responsable del *International Assessment of Educational Progress* (IAEP) y del *Programa National Assessment of Educational Progress* (NAEP)

Podemos atribuirle a la educación comparada, cuyos orígenes datan del siglo XIX, la idea no sólo de comparar sistemas educativos, sino también el aprendizaje de los alumnos (Velloso y Pedró, 1991)

Posteriormente, a raíz de las reformas educativas impulsadas en América Latina a partir de la década de los 90's y en el marco del Proyecto Principal de Educación en América Latina y el Caribe en el periodo 1993-1996 (véase UNESCO 2001),<sup>17</sup> surgen los primeros estudios comparados de evaluación, auspiciados por Organismos Internacionales como la UNESCO y la Organización para la Cooperación y el Desarrollo Académico OCDE. La primera experiencia de evaluación en el contexto latinoamericano la realiza el *Laboratorio Latinoamericano de Evaluación de la Calidad LLECE*, creado en 1994 por la Oficina Regional de Educación para América Latina y el Caribe (OREALC) quien en 1995 aplica el *Primer Estudio Internacional Comparativo sobre Lenguaje, Matemáticas y Factores Asociados para alumnos de Tercer y Cuarto grado* de la Educación Básica, aplicado a un número aproximado de 48,688

En el plano internacional el más reciente sistema de pruebas, el *Programme for International Student Assessment*, mejor conocido como PISA (Programa de Evaluación Internacional de estudiantes), a cargo de la OCDE desde el año 2000 y hasta la fecha durante cada tres años, es quien se ha encargado de realizar estudios para medir el rendimiento escolar de estudiantes

---

<sup>15</sup> Cabe mencionar que la *International Association for the Evaluation of Educational Achievement* (Asociación Internacional para la Evaluación del Logro Educativo) IEA, no nació como Agencia Internacional de Evaluación, "los orígenes de esta asociación se remontan a inicios de los cincuenta, cuando pequeños grupos de profesores iniciaron conversaciones informales en torno al entonces Instituto de Educación de la UNESCO. El interés común que reunía a las personas en esa época era el área de educación comparada, pero con el tiempo el énfasis vanó hacia las técnicas aplicadas de evaluación educacional, que permitiera hacer comparaciones válidas entre naciones. La IEA, se formalizó como asociación en 1964 y ha continuado funcionando hasta el presente" (Frank, 1997, p 117). De ahí que podamos atribuirle a una de las fases de educación comparada cuyos orígenes datan del siglo XIX, fase la idea no sólo de comparar sistemas educativos, sino también el aprendizaje de los alumnos (Velloso y Pedró, 1991)

<sup>16</sup> Su sitio en Internet lo podemos encontrar en <http://www.ets.org>

<sup>17</sup> Véase UNESCO, 2001

de 15 años en tres áreas básicas: Comprensión lectora, matemáticas y ciencias, con el enfoque de "habilidades y conocimientos para la vida" <sup>18</sup>

El cuadro 3 presenta una síntesis de la evolución de las acciones generadas alrededor de los test

Cuadro 3

Desarrollo de los exámenes a gran escala

Época	Etapas	Principales acciones
1900-1915	Precursora	Elaboración de las primeras escalas para medir el rendimiento escolar
1915-1930	Auge	Elaboración y aplicación de una diversidad de baterías de test bajo los indicios de una normalización
1930-1940	Cuestionamiento	Necesidad de cuidar el uso que se les daba
1950-2000	Internacionalización	Primeros estudios de evaluación comparada a nivel internacional

El desarrollo de los tests y exámenes a gran escala y la proliferación que éstos han tenido, especialmente a partir de 1990, con miras a incrementar y comprar la calidad de los sistemas educativos, lleva a repasar en la vigencia de las primeras intenciones con respecto al uso de los tests estandarizados y por otra, al olvido de los señalamientos sobre los efectos perjudiciales de los mismos, ocurridos, ya desde los años cuarentas

El nacimiento y adopción de la medición psicológica y educativa como sinónimo de evaluación y eficiencia educativa es un movimiento tanto sociológico como psicológico, de herencia francesa y estadounidense, que desde sus inicios ha respondido a las situaciones históricas, sociales y económicas de cada época en particular. En el momento actual no es la excepción, pues los tiempos de globalización que estamos viviendo, son tiempos en donde la certificación para la acreditación a través de la medición del rendimiento educativo, se han convertido en la meta educativa de nuestra era, y los exámenes masivos tanto a nivel nacional como internacional han comercializado a la educación, o como diría Meirieu (2004)<sup>19</sup> han estado sujetos a la lógica mercantil

<sup>18</sup> Un análisis de qué y cómo evalúa PISA y su relación con las pruebas nacionales lo podemos encontrar en Padilla, 2005

<sup>19</sup> Autor francés, que desde un planteamiento didáctico, propone una reflexión en torno a la función y sentido que desde una perspectiva pedagógica, debe responder tanto la escuela como el docente en la era actual

El aprendizaje es considerado un producto que se puede medir. En ningún momento se habla de conocer o retroalimentar a través de exámenes. El sentido didáctico que en un momento existió en los exámenes escolares, se pierde por completo, en el momento de aparición y proliferación de los exámenes a gran escala para medir y comparar el rendimiento escolar.

De ahí la necesidad de reconceptualizar a la evaluación del aprendizaje como una actividad que posibilite más que la medición de un logro o resultado, la comprensión del proceso enseñanza-aprendizaje (Díaz Barriga, 1997, 2005). El aprendizaje de los alumnos como objeto de evaluación, y las relaciones que en éste se gestan, es un proceso complejo que demanda un abordaje más pedagógico, que técnico de la evaluación.

Sin embargo las corrientes de pensamiento y enfoques teóricos-metodológicos desde los cuales se ha venido abordando el estudio de los exámenes a gran escala indican una tendencia centrada en la medición, en la rendición de cuentas y en la diferenciación social, que revela la escasez de un abordaje más conceptual y pedagógico para los exámenes a gran escala.

## **1.2 El estudio de los exámenes a gran escala. Un acercamiento al estado de la cuestión**

### **1.2.1. Perspectivas teórico-metodológicas.**

Una revisión de las perspectivas teórico-metodológicas desde las cuales se ha abordado el estudio de los exámenes a gran escala permite identificar tres corrientes de pensamiento:

#### **1) La de la teoría del test**

Llamamos teoría de la medida o teoría del test a los modelos para determinar las puntuaciones de una prueba a gran escala, en donde se construyen matrices de datos de las respuestas que una muestra de sujetos ha dado a cada uno de los ítems, con el fin de estimar el nivel en que poseen los individuos, las características que mide el test y los valores escalonarios de los ítems (parámetros), de ahí que la teoría del test o de la medición, implica cierta manera de hacer la prueba. Para medir o estimar las características latentes de los

sujetos, éstas se relacionan con la actuación observable en una prueba utilizando funciones matemáticas para su descripción

La literatura existente en torno a la teoría de la medida reporta enfoques, modelos, investigaciones y tendencias que dan cuenta del desarrollo obtenido en este campo

### **I. Enfoques:**

Hasta el momento, dos son los enfoques que se reconocen en la teoría del test para el desarrollo de pruebas a gran escala

#### **➤ La normalización de los exámenes**

Entendemos como normas, los marcos de referencia para interpretar las puntuaciones de una prueba. Se trata de estadísticas que describen los resultados de una prueba aplicada a un grupo determinado de alumnos (García Cortés, 1979). Con base en la campana de Gauss, se elabora una distribución normal de la frecuencia en las respuestas correctas o calificaciones de los alumnos en un grupo, lo que determina su posición dentro de éste. Ésta distribución normal de puntajes es la base para el análisis de reactivos y la validación estadística de una prueba tipificada<sup>20</sup>, así como para determinar el nivel de dificultad de los reactivos porque mediante la distribución normal que se hace de los puntajes de una prueba se eliminan los ítems demasiado fáciles o difíciles según la distribución y se dejan aquellos que se encuentran en el promedio

Este tipo de enfoque ha recibido varias críticas y cuestionamientos a las que ya hemos hecho referencia (véase Díaz Barriga, 1997), pues la calificación que obtenga el sujeto en una prueba dependerá del lugar que ocupe en la distribución normal. Lo que importa no es lo que el individuo domina, sino más bien lo que los otros pueden hacer. Y su calificación, dependerá del lugar que ocupe en ésta distribución. El valor y juicio de cada individuo se da manera porcentual. Ejemplo: Pedro, está dentro del 20% de los mejores. De ahí que éste sea uno de los enfoques más utilizados para las pruebas cuyo fin es la selección

---

<sup>20</sup> Llamamos prueba tipificada a las pruebas sujetas a un control experimental a las que también se les conoce como estandarizadas o normalizadas, en función de que existe un procedimiento fijo para elaborarlas e interpretarlas

### ➤ El enfoque criterial

A partir de los años sesenta, con Glaser, surge un enfoque distinto para la validación estadística de una prueba, al que se le ha denominado como criterial

Este enfoque cambia los estándares para validar los reactivos de una prueba y en vez de que los estándares sean fijados con base en un grupo de referencia, es decir, en lo que otros pueden realizar, se establecen en relación a un objetivo considerado como criterio para medir la ejecución de los estudiantes. La calificación de un estudiante se determina así, por la conducta, actuación o desempeño que muestre no en relación a los otros, sino en relación a un criterio o dominio establecido

Aquí los estándares están dados por criterios de conocimiento previamente establecidos. Se trata de un enfoque que propone medir el desempeño individual en relación con dominios específicos establecidos de antemano en un programa de estudios. La taxonomía creada por Bloom en los años cincuentas (véase Bloom, 1971) en relación a distintos niveles de comportamiento o conocimiento con base en los cuales se evaluaba el aprendizaje, es un antecedente importante de este enfoque

Tenemos entonces que el enfoque de evaluación criterial, surge como respuesta a la evaluación basada en la norma o en la curva normal de frecuencias, tiene como característica principal apreciar el grado en el que se han cubierto los objetivos de aprendizaje, en función de las realizaciones de cada alumno, sin compararlo con la de los demás compañeros. El modelo normativo es una evaluación basada en las diferencias individuales de rendimiento, asumiendo la distribución normal. En la evaluación criterial lo que interesa es la tarea misma, sin importar la comparación entre el rendimiento del que se examina y el de los demás, lo que interesa es la comparación con los criterios u objetivos establecidos (Gómez, 1990)

García Cortés (1979), establece una clara distinción de cómo se determinan los estándares en cada uno de éstos enfoques entre ambos enfoques

Cuadro 4

**Estándares referidos a una norma y estándares referidos a un criterio**

Referidos a una norma	Referidos a un criterio
<ul style="list-style-type: none"> <li>• Puntos de referencia relativos (Conforme al promedio obtenido por un grupo de estudiantes)</li> <li>• Los juicios se establecen con base en el promedio "inferior a un promedio", "promedio", "superior al promedio"</li> <li>• No sirve para identificar el grado de dominio de un objetivo por parte de los estudiantes</li> <li>• Los reactivos de una prueba se evalúan con referencia a los estudiantes</li> <li>• Los resultados de las pruebas se interpretan con base en la posición que ocupa un estudiante en relación con las calificaciones de los demás</li> </ul>	<ul style="list-style-type: none"> <li>• Puntos de referencia fijos y claramente determinados</li> <li>• No se describe el promedio</li> <li>• Identifica a los estudiantes que han dominado o alcanzado determinado objetivo</li> <li>• Los reactivos de una prueba se evalúan con base en objetivos</li> <li>• Los resultados de las pruebas se interpretan con relación a la posición que ocupa un estudiante en relación con el plan de estudios</li> </ul>

\*Fuente Tomado de García Cortés (1979)

Éste es el enfoque que sustenta a las actuales pruebas masivas que se aplican tanto a nivel nacional como internacional. Sin embargo cabe destacar que las pruebas referidas a un criterio, están orientadas a medir el nivel alcanzado por un estudiante con base en ciertos objetivos de aprendizaje preestablecidos, que aunque se expresan en desempeños, no dejan de ser una medida conductual desagregada en conocimientos o en conductas, que bajo los criterios de validez y confiabilidad, caen en la trampa de la objetividad del aprendizaje, al ser un procedimiento mediante el cual se trata de evidenciar el dominio de un conocimiento o conducta. Además la idea de medir y comparar a los estudiantes y asignarles una posición, sea con base en los demás o con base en un criterio, es un criterio que prevalece.

**ii. Modelos estadísticos para validación de un examen a gran escala:**

La calibración de los reactivos de una prueba se realiza con base en dos modelos teóricos con bases estadísticas.

▪ **Teoría clásica del test**

Se trata de un modelo lineal desarrollado por Spearman a principios de siglo, en donde el índice de dificultad y discriminación para validar un reactivo está dado por el número de estudiantes que contestaron adecuadamente la prueba

Dentro de la teoría clásica del test las puntuaciones son asignadas con un único valor a cada uno de los ítems, valor que generalmente es de 1, independientemente de su grado de dificultad. Si en una prueba de 35 reactivos una estudiante contesta 15 adecuadamente aunque éstos sean los más fáciles y otro estudiante contesta los 15 más difíciles, los dos obtendrán un puntaje de 15 y su calificación se determinará de acuerdo al promedio del grupo

Este es el modelo que ha prevalecido durante gran parte del siglo XX y en el que se basan los planteamientos y desarrollos realizados por autores como Wood (1971) y Anastasi y Urbina (1998), Thorndike y Hagen (1978) y García Cortés (1979) para elaborar y calificar una prueba estandarizada

Sin embargo este modelo presenta el problema de que la dificultad y discriminación<sup>21</sup> de cada ítem están en función de los sujetos a quienes se les aplica la prueba. Es decir que un ítem resulta fácil o difícil en función de las pericias de un grupo de alumnos y un mismo ítem puede tener distinto índice de dificultad en función de la muestra utilizada. Lo que induce a la búsqueda de nuevos desarrollos en este campo, como el de la Teoría de Respuesta al ítem (TRI)

---

<sup>21</sup> Procedimientos estadísticos para la validación de un reactivo que de acuerdo con la teoría del test, el primero se refiere al porcentaje que da cuenta de la proporción inversa del número de estudiantes que resuelven de manera correcta cada reactivo, mientras que el segundo expresa la capacidad que tiene cada reactivo para discriminar entre alumnos que tienen un mejor rendimiento de los que no lo tienen. El presupuesto estadístico que subyace en ambos índices se encuentra estrechamente relacionado con la llamada curva normal de frecuencias establecida por Gauss



▪ ***Teoría de Respuesta al Ítem (TRI)***

Se conoce como Teoría de Respuesta al Ítem al modelo desarrollado por Lord a partir de 1952, que plantea un modelo un tanto diferente para la calibración de los reactivos de un examen

Sin embargo, conviene aclarar que no se trata de un modelo que se contraponga a la TC, sino más bien que la complementa y que presenta un mayor desarrollo de la estadística y un manejo computarizado de los datos. Es decir, se trata de un modelo matemático más desarrollado para la generación de puntajes en pruebas estandarizadas, cuya mayor aplicación comenzó hasta la década de los ochenta y noventa. Es decir, a medio siglo de haber sido creada

A diferencia de la Teoría Clásica, en la Teoría de respuesta al Ítem, como su nombre lo indica, la unidad de análisis básica es el ítem y no el test. Su base radica en la presunción de que existe en cada individuo una capacidad relativa a la que la prueba evalúa, al que se le conoce como "rasgo latente", lo que determina la capacidad de que el sujeto conteste acertadamente los reactivos con base en su dificultad

Los puntajes se calculan mediante modelos matemáticos complejos y software específicos de procesamiento. Éstos generalmente se expresan en una escala de 500 puntos, pero a diferencia de la TC, no tiene un puntaje máximo determinado por el total de reactivos de la prueba, ni tampoco considera el cero absoluto. En el modelo de TRI se centra la escala en el promedio de habilidad de la población. Por ejemplo si un alumno obtiene 500 puntos significa que está en la media de la población. Esta escala representa tanto la capacidad estimada de los individuos, como la dificultad de las preguntas de la prueba, por lo que el puntaje de cada pregunta se establece sólo en función de su dificultad. Una diferencia importante de la TRI, en relación a la TC, es que no todas las preguntas tienen el mismo peso, lo que quiere decir que aunque dos alumnos contesten el mismo número de preguntas, el que responda preguntas más complejas, obtendrá un mayor puntaje

Esta invarianza de parámetros constituye una de las características principales de la TRI y permite mediciones invariantes respecto al instrumento utilizado y a los sujetos implicados, lo

que ayuda a estimar el rendimiento de los individuos de manera independiente a la prueba o al grupo en donde se aplicó. Esto constituye una de sus mayores ventajas, porque no todos los alumnos tienen que presentar la misma prueba, si existe un grupo de preguntas en común, el puntaje de los alumnos puede estimarse en la misma escala

Uno de los primeros modelos que da respuesta a esta invarianza de parámetros es el Modelo de Rasch, surgido en la segunda mitad del siglo XX con el cual la TRI adquiere una rápida expansión. Se trata de un modelo desarrollado por el matemático Danés Georg Rasch, que establece la probabilidad de respuesta de una persona ante un estímulo dado, en términos de la diferencia entre la medida del rasgo de la persona y la medida del estímulo utilizado (Tristán, 1998). Con base en los diversos desempeños del grupo piloto al que se aplica el instrumento, se integra una planilla electrónica a la que se le conoce como *Escalograma de Guttman*, considerando los resultados de los estudiantes en cada reactivo. Para lo cual se toma como base, los alumnos que obtuvieron el más alto puntaje y se señalan en la planilla las respuestas correctas (1) e incorrectas (0) que tuvieron los alumnos en cada pregunta. El desempeño de cada estudiante se compara contra sí mismo. Se puede dar el caso de que un estudiante de bajo desempeño, pueda responder reactivos con mayor grado de dificultad.

Este escalograma es la base para construir lo que se conoce como *Curva Característica del Item* (curva en forma de S) que relaciona de manera funcional la medida obtenida por cada persona y la probabilidad que tienen para responder a un reactivo. A partir de una pendiente que se construye con las dos unidades, se fijan los valores esperados de respuesta para cada examen. Mientras que en la Teoría Clásica, la base es la *Campaña de Gauss*, que presenta un manejo de intervalos con base en la estadística descriptiva (media, mediana y moda), con la que se formula una curva de distribución normal de frecuencias.

zz Muñiz (1997), presenta la evolución cronológica de dicho modelo, la cual tiene sus antecedentes más remotos en 1905, cuando Binet y Simon, anticipan con sus gráficos el concepto de curva característica. Posteriormente Thurstone, elabora curvas semejantes a las curvas características. El trabajo de Lord "Una teoría sobre las puntuaciones de los test", publicado en 1952, marca el inicio de la Teoría de Respuesta al Item. Pero es hasta 1990, con Hambleton, que se elaboran los primeros manuales para la aplicación de este modelo.

Podemos decir, que la aplicación de la TRI en la elaboración de pruebas, posibilita la construcción de instrumentos equivalentes y permite analizar niveles de desempeño

El empleo de este modelo en pruebas a gran escala es reciente, sobre todo en el ámbito nacional. Sin embargo existen algunos estudios realizados al respecto, como el de Castañeda (2002a), que explora la utilidad de los análisis clásicos y de Rasch en la determinación de la consistencia interna de subpruebas cortas de homogeneidad alta

Así, la Teoría Clásica de los Tests, (TC) y la Teoría de Respuesta al Ítem (TRI), constituyen los fundamentos básicos de la Teoría de los tests, que marcan los desarrollos obtenidos en el campo de la medición psicológica

Cabe mencionar que no se trata de modelos distintos, finalmente se trata de modelos estadísticos para la validación de un reactivo en donde los elementos básicos para analizar un test están presentes en ambos reactivos. Lo que varía son los criterios estadísticos utilizados. El Cuadro 5 muestra las principales diferencias encontradas en ambos modelos

Cuadro 5

**Modelos estadísticos para la validación de un examen a gran escala**

Criterios	Teoría del test	
	Teoría Clásica	Teoría de Respuesta al Ítem
Unidad de análisis	Test	Ítem
Determinación de puntaje	Único valor independientemente de su grado de dificultad	En función de su dificultad No todas las preguntas tienen el mismo puntaje
Validación del reactivo	Índice de dificultad y de discriminación	Índice de dificultad

Sin embargo pese a las diferencias en criterios estadísticos encontradas en ambos modelos, tenemos que la base para la construcción y validación de una prueba a gran escala sigue siendo un modelo matemático<sup>23</sup>. Esto da cuenta de que el test, ha sido conceptualizado como

<sup>23</sup> La incorporación de la teoría del test al campo de la evaluación, en miras a cientificar el acto evaluativo, ha ocasionado que ésta se centre en la medición. Si consideramos que medir es asignar números a objetos o acontecimientos de acuerdo a ciertas leyes, comparándolo con una escala o unidad de medida, en función del instrumento y las propiedades del objeto a

---

una situación experimental, estandarizada con fines de comparación estadística para ubicar a los sujetos examinados en una clasificación o tipología (Benedito, 1986)

La finalidad del test es medir y medir comparar. La incursión de la teoría de la medición al campo de la evaluación, en especial de la evaluación del aprendizaje, concibe a éste, como un objeto susceptible de ser medido y a los seres humanos como sujetos de clasificación.

## **2) La rendición de cuentas como parte de una política administrativa**

Como práctica sistemática racional, la evaluación-medición que se realiza de los sistemas educativos está encaminada a la operación, administración y control de los procesos a partir de objetivos previamente establecidos. La evaluación-medición de la acción educativa es concebida entonces, como el abordaje sistemático, racional y organizado del proceso enseñanza-aprendizaje, para la obtención de resultados específicos, potencialmente repetibles y predecibles (ILCE, 1993).

Desde esta concepción, la referencia a problemas prácticos que se pueden resolver, lleva al problema de costo-eficiencia.

El enfoque adoptado por Chadwick, (s/f) desde lo que se llamó el movimiento de la tecnología educativa, es quizá el antecedente de lo que hoy se conoce como evaluación para la toma de decisiones o para la rendición de cuentas.

Stufflebeam y Shinkfield (1987), exponen que el propósito de una evaluación es la utilidad. Su principal objetivo es aumentar la efectividad en la administración del programa, con lo cual la evaluación se convierte en una acción pragmática que adquiere un sentido funcional.

La discusión sobre los sistemas de información para la toma de decisiones, inicia, desde los años ochenta. En el plano latinoamericano Rama (1994), analiza las repercusiones que para los Ministerios de Educación en América Latina, conlleva el que la necesidad de información

---

medir, la validez de una medición, dependerá de las propiedades del objeto a medir y de las del número, lo que está determinado por ciencias como las matemáticas (Benedito, 1986)

para la toma de decisiones sea parte de una gestión moderna de la política educativa y las inequidades que con esto se genera.

Otro planteamiento interesante lo realiza Schiefelbein (1995) quien expone que en la investigación e información para la toma de decisiones, se diseñan y aplican nuevos modelos de administración, planificación y gestión de la educación, a partir de lo cual surgen distintas fuentes de financiamiento. Sin embargo los cambios importantes en la calidad del aprendizaje y en la formación docente, no se han generado.

Así, encontramos que en la obstinación por la aplicación de exámenes a gran escala para la examinación masiva del aprendizaje escolar, como política para la toma de decisiones, prevalece un enfoque gestivo, centrado en la administración y control del rendimiento escolar. Ante lo cual, el examen, de ser una herramienta del docente para revisar el aprendizaje de sus alumnos, pasa a convertirse en un instrumento contra el docente, por las consecuencias que los resultados obtenidos generan en relación a su trabajo, valoración y seguridad o inseguridad profesional. Ante lo cual los exámenes a gran escala, se convierten y asumen como un poderoso y temible instrumento de regulación, más que de conocimiento y ayuda a la tarea docente, lo que a su vez los convierte en instrumentos de comparación y discriminación.

### **3) El test como legitimador de desigualdades sociales**

Si consideramos a los exámenes a gran escala como instrumentos que nacen dentro de la sociedad industrial para dar respuesta a las necesidades socioculturales y educativas de la época, no podemos desconocer el carácter social de esta actividad y el papel que la medición estandarizada, asume en la diferenciación social.

Es conveniente recordar que las primeras mediciones de tipo sensorial, psicofísico y biológico, se realizaron con base en las diferencias existentes, primero entre las especies y luego entre los individuos. El concepto de adaptación científica de Darwin, donde al igual que en el medio natural, los más aptos son los que sobreviven. Y el concepto de desigualdad natural de Francis Galton, quien se ubica como el primero en construir y aplicar test a gran escala, da lugar a que se considere a unos individuos "más aptos y triunfantes" que a otros. Lo que lleva también a una especie de "selección natural" entre éstos. Una de las primeras funciones de los

tests mentales, era realizar sino una "selección" natural, si una selección social, con base en las "capacidades mentales y aptitudes" de los individuos (Benedito, 1986)

"La necesidad de comparar los resultados obtenidos con diferentes individuos hizo que se extendieran los esfuerzos para utilizar la medida en el campo de la psicología y que comenzaran a adaptarse algunas técnicas estadísticas" (Székely, 1966, p 4), donde la comparación estadística, en cuanto al comportamiento de unos sujetos frente a otros, es lo que caracteriza y diferencia a los tests de los exámenes escolares. La finalidad de los tests es la estandarización y comparación a través de la medición. Hablar de estándar es hablar de lo que se ajusta a un modelo o patrón único, que refiere al proceso estadístico mediante el cual se establecen las normas de los puntajes establecidos para un test. Se trata de una situación experimental, fruto de los trabajos en el laboratorio para el estudio de los fenómenos psicológicos, centrada en la observación de los hechos y la conducta (Benedito, 1986, p 180)

La diferenciación social ocasionada por los tests, podemos ubicarla desde tres ámbitos

#### **i. En la división social del trabajo**

Desde que en Francia nacen los primeros test para medir la inteligencia, con Binet y Simon, hecho al que ya hemos hecho referencia, la polémica sobre lo que ocasiona la medición del Coeficiente Intelectual en la reproducción de clases sociales y división del trabajo, ha sido abordada por autores como Tort (1974), para quien el acto de medir la inteligencia, es un dispositivo ideológico del ejercicio escolar, que tiene sus raíces en la división capitalista del trabajo. Y en donde el test es la manera de justificar "científicamente" la división que la escuela produce. Dice Tort (1979) "La medición de la inteligencia, es la medición de la desigualdad de los hombres" (p 2). La supuesta objetividad de los tests, queda en entredicho cuando plantea:

"La situación de un test no es neutra. Es una situación de clase que favorece desigualmente el éxito de los niños según su origen social. Las preguntas que se formulan no tienen en absoluto el mismo sentido para las distintas clases sociales" (p 2)

La homogenización que los exámenes a gran escala adquieren a partir de su estandarización y aplicación masiva a un gran número de estudiantes, no consideran la heterogeneidad y

diversidad educativa y social existente. Lo que los convierte en instrumentos de desigualdad social.

**ii. En su significación ideológica como instrumentos de exclusión, para la conservación política de una sociedad.**

El carácter hereditario de la inteligencia, en donde, según planteamientos de Terman,<sup>24</sup> los niños con Coeficiente Intelectual (CI) bajo en los tests de inteligencia, deberían ser separados de los que obtuvieran un CI alto. Es así como la medición estandarizada de esta cualidad humana, adquiere un carácter determinista y racial.

Con base en lo publicado por Terman (1916), sobre las bondades del test de Binet, en donde expresa que dado el carácter hereditario de la inteligencia, los niños que muestren diferencias significativas, deberán ser segregados en clases especiales y convertirlos en trabajadores eficientes y a las personas detectadas con este problema, no se les debería permitir reproducirse, en varias ciudades norteamericanas se aprobaron leyes que impedían la procreación e inmigración de débiles mentales a Estados Unidos. Esto es refutado por Kamin (1983), quien plantea cómo el uso de los tests, obedecía a intereses políticos e ideológicos, pues eran empleados para justificar leyes que impedían la procreación y migración de "deficientes mentales" en la sociedad americana.

**iii. Como promotor de diferencias sociales.**

En textos de Snyder (1972) y Salvat (1978), podemos identificar, como la escuela con sus "calificaciones" y certificados puede convertirse en una reveladora de desigualdades sociales. Como fruto de una economía liberal, dice Salvat (1978) "El fracaso escolar no es, en efecto, más que un producto de nuestro sistema escolar y una expresión de las fallas de nuestra sociedad" (p. 9).

---

<sup>24</sup> Terman fue a quien se le atribuye a difusión de los tests de Stanford y Binet en Estados Unidos.

Desde que los sistemas educativos adoptaron en el siglo XIX la calificación del examen para aprobar o reprobado a un alumno, es decir para certificar un conocimiento, éstas se han convertido en determinantes de oportunidades sociales y educativas

Tenemos entonces que los sistemas de exámenes con sus escalas de medición y puntuaciones obtenidas, sea conforme a la Teoría Clásica o con la Teoría de Respuesta al ítem, conlleva a una competencia y a una clasificación y etiquetación de escuelas y estudiantes con calificativos como "aprobado o reprobado, o bien de "los mejores y los peores"<sup>25</sup> Esto produce una lucha por parte de alumnos para ocupar un lugar y sobrevivir. Con lo cual el fin de la educación establecido por Herbart en el siglo XIX, de una educación del carácter y la voluntad, que responda a la multiplicidad del interés de un individuo para educarlo en la voluntad<sup>26</sup> se corrompe

Estamos ante tres enfoques bajo los cuales se investiga y aborda el tema de las pruebas a gran escala para medir el rendimiento escolar. El de la medición, el de la rendición de cuentas y el crítico-social, desde los cuales, podemos conceptualizar a los exámenes a gran escala aplicados en forma masiva, como instrumentos de medición (con todo lo que esto implica), de control y discriminación

El estudio de los exámenes a gran escala desde sus diferentes perspectivas de análisis, lleva a identificar también algunas tendencias y aplicaciones

### 1.2.2 Investigaciones recientes: Tendencias y aplicaciones :

Una revisión rápida a la Base Electrónica de los Recursos Informativos del Centro de información (ERIC)<sup>27</sup> reporta un total de 12, 716 artículos sobre test,<sup>28</sup> lo que revela una amplia

<sup>25</sup> Algunas notas de prensa al respecto, que ilustran esta aseveración "Dan palomita a secundarias capitalinas" Destacan escuelas donde estudiaron jóvenes que alcanzaron un lugar destacado en el DF y los 22 municipios corribados con el examen único" (Romano, 2003 *Reforma*, 13 agosto) "Publica CENEVAL Cuadro de honor" (Romano, 2003, 13 de agosto, 2003)

<sup>26</sup> Dice Herbart (1806-1983) "Sólo la voluntad misma del hombre futuro, y, por consiguiente, la surra de las exigencias que tendrá para consigo mismo por y con esta voluntad, constituye para el educador el objeto de su berevolencia" (p. 46)

<sup>27</sup> Su sitio en internet <http://www.eric.ed.gov/ERICWebPortal/Home.portal>

<sup>28</sup> Datos tomados hasta septiembre de 2008 [http://www.eric.ed.gov/ERICWebPortal/Home.portal?\\_nfpb=true&ERICExtSearch\\_SearchValue\\_0=TEST&ERICExtSearch\\_SearchType\\_0=i&pageLabel=ERICSearchResult&newSearch=true&nd=1222630064734&searchType=keyword](http://www.eric.ed.gov/ERICWebPortal/Home.portal?_nfpb=true&ERICExtSearch_SearchValue_0=TEST&ERICExtSearch_SearchType_0=i&pageLabel=ERICSearchResult&newSearch=true&nd=1222630064734&searchType=keyword) consultado en septiembre del 2008



literatura existente en torno a los tests como situación estandarizada. Simplemente de 1998 a 2001, encontramos 92 artículos, que permiten ver la evolución teórica y metodológica de estos instrumentos y la enorme y variada producción internacional que hay al respecto. Las temáticas van desde como validar un test, la teoría del test, el análisis de ítems, los tests asistidos por computadora, la construcción de un test, el uso de los tests y el quehacer de asociaciones internacionales de test como la National Assessment of Educational Progress (NAEP)<sup>29</sup>. Además de revistas electrónicas especializadas sobre investigación y evaluación a nivel internacional como la *Practical Assessment Research and Evaluation*<sup>30</sup> en donde de 1998 al 2008, se han publicado 39 artículos sobre test.

En el plano nacional, localizamos algunas investigaciones recientes, en el ámbito de la medición psicológica y educativa que revelan algunas aplicaciones y tendencias.

*El estado del arte sobre la evaluación y el fomento del desarrollo intelectual en la enseñanza de las ciencias, artes y técnicas* (Castañeda, Lugo, Pineda y Romero, 1998), es una investigación con enfoque funcional, cuyo objeto de estudio reside en identificar tendencias dominantes y emergentes de las aproximaciones teórico-metodológicas sobre el fomento y la evaluación del desarrollo académico cognitivo en la enseñanza y el aprendizaje.

Se trata de un modelo cualitativo que consideró categorías de análisis con diferentes niveles de inclusividad que comprenden tanto componentes de evaluación, como de desarrollo cognitivo. Las tendencias identificadas como emergentes residen en:

- ✓ El análisis cognitivo de tareas
- ✓ El análisis de la pericia, (comportamiento experto)
- ✓ Las pruebas referidas a criterio
- ✓ Evaluación de competencias
- ✓ Evaluación del desarrollo cognitivo

La tendencia actual en materia de medición, es la disminución de pruebas con base en la norma y el incremento de pruebas referidas a criterios, así como la incursión de la psicometría cognitiva, basada en el estudio de las pericias y el análisis cognitivo de tareas.

---

<sup>29</sup> Su sitio en Internet es <http://www.naep.org/>

<sup>30</sup> Su sitio en Internet <http://pareonline.net/getpage.asp?wh=8&abt=test>

La utilidad del Análisis Cognitivo de Tareas, documentada desde los años sesenta, ha tenido algunas aplicaciones en México. Éstas han sido realizadas por Castañeda (1998), para realizar *Modelos Multicomponentiales de Medición de Resultados de Aprendizaje*. Este modelo constituye un auxiliar en la construcción de exámenes de resultados de aprendizaje, bajo diseño de modelos de observación y medición y está sustentado en los avances generados en materia de cognición aplicada al campo educativo. Ha sido desarrollado por el Laboratorio de Desarrollo Cognitivo e Innovación Tecnológica de la Facultad de Psicología de la UNAM, con una doble finalidad. Por una parte apoyar la identificación, representación y organización sistemática de conocimientos y habilidades y por otra, favorecer el diseño de tareas e indicadores que supuestamente actúan en esos conocimientos y habilidades. Lo que requiere establecer descripciones claras del conocimiento tanto conceptual como procedimental y estratégico, que se realiza mediante una desagregación de las tareas implicadas en los diferentes niveles de complejidad implícitos en los diferentes tipos de conocimiento (Castañeda, 2002b).

También podemos encontrar en Castañeda (2005), una investigación sobre la evaluación del aprendizaje en educación superior con enfoque cognitivo y funcional, que presenta una caracterización de tareas y reactivos en relación al atributo a medir e identifica fuentes de contenido que generan error y afectan los puntajes. Este trabajo se realizó con el objeto de identificar la magnitud con la que cuatro fuentes de contenido incluidas en los ítems co-ocurrían con el puntaje global de dificultad a priori del reactivo y de esta manera establecer si habría efectos diferenciales del tipo de conocimiento incluido en esas fuentes sobre la dificultad empírica de los ítems. Se tomó como referencia el análisis funcional y el análisis cognitivo de tareas, a partir de un análisis cuantitativo de las fuentes de contenido incluidas en los ítems (demandas cognitivas conceptuales y de andamiaje).

La investigación en materia de evaluación cognitiva (véase Wittrock y Baker, 1998), permite identificar el uso de procesos y estrategias cognitivas y motivacionales en la elaboración y desarrollo de pruebas para medir el aprendizaje. Las investigaciones sobre cognición de Glaser, Sternberg, Weinstein, Combs, Collins y Romberg, Baker, Freeman y Claiton, que presentan un interés por estudiar la implicación de la atención, motivación, comprensión, estrategias de aprendizaje y cognición en la aplicación de pruebas, son muestra de ello.

La revolución tecnológica e informática, ha incursionado fuertemente también en los sistemas de exámenes, sobre todo en los aplicados de forma masiva, tanto para su calificación, como es el caso de los exámenes aplicados en la Facultad de medicina de la UNAM en el año de 1962, como para su aplicación y elaboración. Hoy en día, es una realidad que los exámenes por computadora, comienzan a sustituir a los exámenes de lápiz y papel y la evaluación asistida por computadora, es una de las tendencias recientes.

La Universidad Autónoma de Baja California, a partir de 1994 desarrolla un Sistema Computarizado de Exámenes, conocido como (SICODEX), versión electrónica de la versión de lápiz y papel del *Examen de Habilidades y Conocimientos Básicos* (EXHCOBA) (Backhoff y Tirado, 1992). Sin embargo se trataba de un examen normativo con propósitos de selección que no proporcionaba información sobre el nivel académico de los estudiantes.

El desarrollo obtenido en materia de informática y la existencia de la Teoría de Respuesta al Ítem, ha posibilitado la realización de versiones automatizadas de pruebas adaptativas. La evaluación adaptativa por computadora, desarrollada por Lord y Weiss en los setenta, constituye una de las tendencias recientes en materia de elaboración de exámenes a gran escala. La idea que subyace en este planteamiento es la aplicación de una prueba con un diseño individualizado para cada uno de los examinados, con el propósito de obtener una estimación más eficiente de las habilidades específicas de todos y cada uno de los miembros del grupo de aplicación.

Un ejemplo de esta modalidad es el Sistema Computarizado de Exámenes (SEA), para administrar el Examen de Ubicación Matemática, (EXUMAT) aplicado desde 1996, por la Universidad Autónoma de Baja California. El SEA contiene siete diferentes tipos de preguntas, que exige a los estudiantes respuestas específicas (véase Backhoff y Rosas, 2002). El desarrollo, validación e implementación de este Sistema para la selección de estudiantes en la Universidad Autónoma de Baja California, constituyó la tesis doctoral de Backhoff (véase Backhoff, 2003).

Si bien, la investigación en torno a exámenes a gran escala ha girado en relación a la medición cognitiva y a la incursión de la informática a la elaboración, aplicación y calificación de exámenes, prevalece la escasez de un tratamiento más conceptual y de construcción teórica del examen como instrumento de evaluación. Existen estudios que inician el debate teórico en

relación al contenido y tipo de ítems de los exámenes estandarizados tanto en el plano internacional como nacional (véase Padilla, 2005 y Díaz Barriga, 2006). En Díaz Barriga (2006), encontramos un análisis de las diferencias que subyacen en cuanto a Teoría del test, concepción de aprendizaje y construcción de reactivos, en dos exámenes masivos: El *Examen Nacional de Ingreso a la Educación Media Superior* (EXANI) y el del *Proyecto Internacional de Evaluación de Estudiantes* (PISA), por sus siglas en inglés. Y en Padilla (2005) un análisis comparativo de los sustentos conceptuales entre pruebas nacionales e internacionales.

Las investigaciones revisadas, en el ámbito de la medición educativa y psicológica, permiten identificar a los autores que representan tres enfoques distintos de la misma corriente:

El interés por la medición cognitiva del aprendizaje, trabajada por (Castañeda, Lugo, Pineda y Romero, 1998, Castañeda, 2002 a y b y 2005), los exámenes a través de sistemas computarizados, a cargo de Backhoff (1998, 2002 y 2003) y un interés por trabajar una teoría del contenido, en exámenes a gran escala tanto nacionales, como internacionales (Díaz Barriga, 2006).

Los desarrollos y tendencias recientes en materia de la investigación sobre exámenes a gran escala, a los que hemos hecho alusión, lleva a una reflexión. A pesar de los distintos enfoques, modelos y aplicaciones que se han dado en materia de test de rendimiento escolar, la premisa sigue siendo la misma: medir el aprendizaje y a los sujetos, sea con base en una norma, sea con base en un criterio o dominio, sea en cuanto a conocimientos, habilidades o estrategias cognitivas, con lápiz o papel o a través de una computadora. Pero el propósito y sustento siempre es el mismo, la medición y cuantificación del aprendizaje dentro de los sistemas educativos.

Una revisión de los modelos existentes para el análisis de reactivos de un examen a gran escala, revela la inclusión de estas tendencias:

### 1.2.3 Modelos de análisis para los exámenes a gran escala. Un acercamiento a lo que se ha hecho.

En materia de análisis de pruebas a gran escala, se pueden identificar tres tipos de modelos

- a *Modelos psicométricos* Consiste en el estudio de los parámetros de los reactivos en función de las propiedades psicométricas de los tests. Se trata de un proceso de selección de ítems a través del cual se decide que ítems son pertinentes y cuáles no, de acuerdo con las medidas arrojadas. Los criterios de análisis dependen del modelo teórico bajo el cual se construyó el test (García y Fidalgo, 2005). Es decir si se siguió el de la Teoría clásica o el de la Teoría de Respuesta al Ítem. Pero en ambos modelos, sobre todo, se toma en cuenta el índice de dificultad y de discriminación de cada reactivo.
- b *Modelos cognitivos* Se refiere a una técnica para identificar mecanismos subyacentes al aprendizaje deseado en grado de complejidad creciente (Castañeda, 2006). Se presenta como una de las tendencias emergentes en la evaluación del aprendizaje y consiste en descomponer una tarea determinada en aquellos componentes considerados como esenciales para llevar a cabo una de las subtareas que lo constituyen. Lo que interesa es conocer con precisión los factores que componen una tarea global y poder identificar en qué paso de la tarea el estudiante puede tener problemas. Este tipo de análisis alude al elemento cognitivo, es decir incluye la identificación de los procesos, las estructuras y las estrategias psicológicas que subyacen a cada subtarea componente. Los criterios que considera para su realización son los siguientes:
- Contexto: se refiere a las condiciones bajo las cuales se realiza la tarea.
  - Contenido: se refiere al análisis del tipo de información que debe ser procesada durante la tarea.
  - Procesos cognitivos, se identifican procesos de entrada, de transformación y de salida.
  - Estrategias: tienen que ver con reglas y procedimientos de planeación para lograr una ejecución óptima.
  - Tareas: criterios: especifica la demanda de las tareas y la forma en que será valorada la ejecución (Castañeda, Lugo, Pineda y Romero, 1998).

c. *Modelos para analizar la traducción* Con el propósito de evaluar la calidad de la traducción de los reactivos de pruebas internacionales como el *Tercer Estudio Internacional de Matemáticas y Ciencias*, (TIMSS) y el *Programa Internacional de Evaluación de Estudiantes* (PISA) Solano-Flores, Contreras-Niño y Backhoff (2006), realizaron un estudio donde desarrollaron una tipología de errores de traducción de reactivos. Para el desarrollo de esta tipología se basaron en un sistema de codificación de errores de traducción que emplea el ATA (American Translators Association, 2003). Los errores de traducción fueron agrupados en siete categorías:

- *Estilo y ortografía*: la traducción del reactivo no es consistente con las convenciones estilísticas y ortográficas utilizadas en textos impresos en el país.
- *Formato*: la información visual del reactivo en español se presenta de manera diferente al reactivo en el idioma original.
- *Convenciones*: la traducción no refleja el estilo convencional de reactivos en el idioma Español o viola principios de redacción de reactivos aceptados universalmente.
- *Gramática y sintaxis*: la estructura gramatical del reactivo no es tan natural como en su escritura original en español o no es adecuada para la edad del estudiante examinado.
- *Información*: la traducción altera la cantidad o el contenido de la información que se proporciona al estudiante para entender las instrucciones o el problema que se le presenta.
- *Semántica*: las ideas expresadas en el reactivo no tienen el mismo significado que en la lengua original, debido a que la traducción es imprecisa (Solano y Backhoff, 2003, p.4).

Los diferentes tipos de análisis de reactivos encontrados demuestran un énfasis en modelos estadísticos y psicológicos. Hasta este momento hay una omisión en tomar en cuenta las aportaciones didácticas al campo de la enseñanza. Contrario a lo que se afirma, que no es factible evaluar algo que no ha sido enseñado, la visión didáctica, que es la que daría cuenta de esta cuestión, se encuentra ausente.

Esto lleva a pensar en la necesidad de construir un modelo didáctico para el análisis de una prueba. Un tipo de análisis que abandone el tratamiento meramente estadístico y psicológico y permita desde un enfoque didáctico identificar los elementos que se ponen en juego en las actividades de evaluación y su relación con el proceso de enseñanza y aprendizaje.

Una revisión de la producción generada en México sobre exámenes a gran escala en los últimos años, permitirá ubicar el impacto social que éstos están teniendo a raíz de su la institucionalización y politización del examen

#### **1.2.4 Producción e institucionalización de los exámenes a gran escala en el escenario nacional (1995-2007)**

Un rastreo y acopio de los documentos publicados en páginas electrónicas especializadas de los Institutos Nacionales de Evaluación<sup>31</sup> permitió caracterizar la producción sobre exámenes masivos en México. Se recuperaron 155 documentos: 12 de la DGEE, 92 del INEE y 51 del CENEVAL los cuales fueron objeto de una clasificación inicial conforme al tipo de publicación.

El Cuadro 6 muestra la distribución cuantitativa de los documentos recabados

---

<sup>31</sup> Los institutos y páginas consultadas a las que nos referimos son: Dirección General de Evaluación Educativa de la SEP (DGEE) <http://www.sep.gob.mx/portal/apps/ed/dge/index.htm>, Centro Nacional para la Evaluación de la Educación Superior (CENEVAL) [www.ceneval.edu.mx](http://www.ceneval.edu.mx) e Instituto Nacional para la Evaluación de la Educación (INEE) <http://www.inee.edu.mx/>. Cabe señalar que el orden en que los señalamos, corresponde a su momento de aparición, siendo la DGEE, cuyos orígenes datan de 1970 la que se ubica como la primera institución responsable de la evaluación de la educación básica a nivel nacional. El segundo instituto conforme su aparición es el CENEVAL, creado en 1994, y el de más reciente creación es el INEE. Nótese la existencia de tres institutos para una misma acción en el sistema educativo, además cabe destacar que dos de ellos, fueron creados en la década de los noventa.

Cuadro6

**Documentos publicados por institutos de evaluación  
1995-2007**

Años de publicación	Institución	Total de documentos publicados	Tipo de publicaciones													
			Informes				Estudios						Discusiones			
			ID	IT	IA	IC	EC	ED	EA	ECC	ET	EX	DD	DT	DA	DC
1999-2004	DGEE	12	2	1	-	-	-	2	-	1	-	-	4	1	1	-
2003-2008	INEE	92	31	28	1	-	-	3	1	-	13	-	7	4	4	-
1995-2007	CENEVAL	51	2	14	-	-	-	8	1	1	16	-	-	8	1	-
		155														

Claves:

ID = Informe descriptivo

IT = Informe técnico

IA= Informe analítico

IC= Informe conceptual

EC= Estudio comparativo

ED= Estudio descriptivo

EA= Estudio analítico

ECC= Estudio conceptual

ET = Estudio técnico

EX= Estudio experimental

DD = Discusión descriptiva

DT = Discusión técnica

DA= Discusión analítica

DC= Discusión conceptual



Un análisis de la información presentada en el Cuadro 6, permite observar como el INEE se ubica como la institución con mayor número de publicaciones sobre pruebas estandarizadas, sin embargo 31 de éstos, es decir el 33.6 % consisten en informes descriptivos y 28 es decir el 30.4 % en informes técnicos.

El CENEVAL por su parte es el que más estudios ha realizado, sin embargo éstos también son de tipo técnico. Lo que habla de un tratamiento meramente técnico en la información sobre exámenes estandarizados. El abordaje y la discusión conceptual y analítica aparece como prácticamente inexistente. Además de que casi en su totalidad los destinatarios de estos documentos son la sociedad y las autoridades.

Las temáticas abordadas en los textos publicados por estos institutos giran básicamente en torno a:

- ✓ Elaboración y aplicación de pruebas nacionales e internacionales
- ✓ Resultados de las pruebas internacionales y nacionales
- ✓ Estudios de factores asociados
- ✓ Reportes técnicos
- ✓ Informes cuantitativos

Esto permite dar cuenta que la producción sobre exámenes a gran escala en México de las últimas décadas, ha sido acaparada por los Institutos Nacionales de Evaluación creados en la última década del siglo XX, quienes se disputan la hegemonía en materia de elaboración de exámenes sin ninguna coordinación entre sí.<sup>32</sup> Institutos Nacionales de Evaluación como el INEE y el CENEVAL, así como la Unidad de Planeación y Evaluación de Políticas Educativas (UPEPE)<sup>33</sup> de la SEP, compiten por el reconocimiento en materia de elaboración de exámenes.

Sin embargo la producción generada por estas instituciones en materia de elaboración de exámenes, consisten en informes destinados a la sociedad para evidenciar deficiencias y

<sup>32</sup> Un recuento de los exámenes tanto nacionales como internacionales que presentan estudiantes de educación básica en México, arroja que, dependiendo del grado en el que se encuentren, aproximadamente son cinco los exámenes a gran escala, que los alumnos de educación básica, presentan cada año: PISA, IDANIS, EXANI-1, EXCALE y ENLACE (independientemente de los estatales e institucionales), lo que ocasiona que gran parte de la acción docente y educativa esté en función de estos exámenes.

<sup>33</sup> Cabe destacar que este nombre de la dependencia encargada de elaborar y aplicar el examen a gran escala con mayor cobertura en México (ENLACE), hace pensar en que los exámenes a gran escala en México, se han vuelto parte de una política educativa, más que de una acción didáctica.

comparar a los “mejores y los peores”, con el fin de buscar una sanción social. El contenido de éstos, se enfoca principalmente a la presentación de resultados y a los procedimientos técnicos que a éstos subyace, lo que denota una ausencia de referentes conceptuales y un epistemicidio en lo que respecta a pruebas estandarizadas.

Su preocupación por “informar” a la sociedad los resultados en términos cuantitativos, deja fuera a uno de los protagonistas principales del acto educativo y de enseñanza: el docente. Siendo éste el único agente que si se le proporciona una información sobre el contenido y principios que subyacen en los exámenes, puede realizar los ajustes necesarios en materia de enseñanza-aprendizaje.

La revisión y análisis sobre la producción de los exámenes a gran escala en México, deja ver lo siguiente:

- La experiencia mexicana en cuanto a test, ha adoptado la mirada norteamericana de medir para eficienciar. Ésta es la línea que se asume en las acciones y políticas en materia de evaluación a nivel nacional.
- Aunque la producción sobre exámenes a gran escala ha ido aumentando y ahora contamos con un mayor número de publicaciones, lo cual representa un avance en este campo. Estas publicaciones conservan un enfoque centrado en la medición y dejan fuera la discusión y conceptualización en torno a la evaluación del aprendizaje a través de exámenes estandarizados.
- La información publicada sobre los exámenes masivos elaborados y aplicados por las diferentes instancias de evaluación en nuestro país ¿ha ayudado en alguna manera a mejorar las condiciones y el aprendizaje de la educación básica en México?

### 1.3 La aplicación de exámenes a gran escala como parte de una política educativa

La producción y aplicación de exámenes a gran escala en nuestro país, ha venido a formar más bien parte de una política educativa que poco o nada tiene que ver con lo educativo

El Programa Sectorial de Educación 2007-2012, (2007) establece como uno de sus primeros objetivos que "la evaluación será un instrumento fundamental en el análisis de la calidad ( ) La evaluación debe contemplarse desde tres dimensiones como ejercicio de rendición de cuentas, como instrumento de difusión de resultados a padres de familia y como sustento del diseño de las políticas públicas"(p 11) Los resultados obtenidos en las pruebas a gran escala tanto nacionales como internacionales, se han convertido en un indicador de calidad educativa y en la meta principal de la educación hoy Las metas e indicadores de calidad para el 2012 son las siguientes

Nombre del indicador	Unidad de medida	Situación en 2006	Meta 2012
Calificación en el examen PISA en las pruebas de matemáticas y comprensión de lectura	Puntaje entre 200 y 800 (800 equivale al mejor rendimiento)	392 (de acuerdo a resultados 2003)	435
Porcentaje de alumnos con un logro académico al menos elemental en la prueba ENLACE (los niveles de la prueba son insuficiente, elemental, bueno y excelente)	Porcentaje de alumnos examinados	Primaria Español=79.3% Matemáticas=79% Secundaria Español=59.3% Matemáticas=38.9%	Primaria Español=82% Matemáticas=83% Secundaria Español=70% Matemáticas=53%

Fuente: SEP, 2007. Programa Sectorial de Educación, p 15

Al parecer la premisa central de la actual política educativa, *es a mayor cantidad de exámenes, mejor calidad de la educación*.<sup>34</sup>

Además, la Unidad de Planeación y Evaluación de Políticas educativas de la SEP, declara la intención de asignar un bono equivalente a cinco salarios mínimos a los profesores, cuyos alumnos hayan demostrado un incremento en los resultados de la prueba ENLACE

“Cuando inicie el ciclo escolar (en agosto) se le dará al profesor un diagnóstico de su grupo, que se construirá con ENLACE en el cual se informarán los parámetros que sus alumnos obtuvieron promedio, porcentaje máximo y mínimo, entre otros. Al final del curso (mayo del siguiente año) se volverá a aplicar ENLACE en la medida en que la foto inicial y final del curso cambien para bien, el profesor tendrá un estímulo como único pago” (Pavón, 2008)

Tal decisión pone en tela de juicio a nuestros estudiantes, a nuestros maestros y a nuestras escuelas. Además de que promueve los vicios ocasionados por la centralización de la evaluación en una calificación.<sup>35</sup>

Lo que da cuenta que el examen pasó de ser un recurso que el docente empleaba para evaluar el aprendizaje de sus alumnos, a un instrumento estandarizado de gestión y verificación curricular, para la toma de decisiones y rendición de cuentas. Esto deja de manifiesto la intención judicial, más que educativa que ha adquirido el examen a gran escala hoy.

Tal pareciera que la apuesta de hoy en materia de evaluación del aprendizaje es *“a mayor cantidad de exámenes, mejor calidad de la educación,”* lo que deja ver una vez más, que el

---

<sup>34</sup> Las cursivas son nuestras

<sup>35</sup> Giner de los Ríos (s/f) refiere como el sistema de calificaciones al que se vinculó el examen desde mediados del siglo XX desvirtuó a la educación, a la docencia y por ende a la pedagogía y mató el placer por aprender en los estudiantes, al centrar y privilegiar la acción de alumnos y maestros en torno al examen. “No hay más que una necesidad, ser aprobado, llevarse la nota, el premio, la plaza.” (Díaz Barriga, 1993<sup>2</sup>, p 74)

elemento central que subyace a la aplicación de exámenes masivos en el país, es carácter político y administrativo, más que educativo

La aproximación al estado en el que se encuentran el campo de los exámenes a gran escala y la manera como éstos han sido conceptualizados y estudiados, permite identificar algunas tendencias que vale la pena destacar

- La búsqueda de la productividad, objetividad y eficiencia en materia de educativa son los principios que originaron y prevalecen en la medición educativa
- La aplicación de exámenes a gran escala como sinónimo de calidad educativa, es una medida más de corte administrativo que educativo
- Una concepción fiscalista y estadística, es la que prevalece en el tipo de informes sobre exámenes a gran escala que se generan en el ámbito nacional
- El aprendizaje continúa siendo conceptualizado como algo que se puede medir
- La intención con la que se inicia la aplicación de exámenes de impulsar una reforma en la enseñanza, ha sido suplantada por la idea de controlar y homogenizar el rendimiento escolar
- La elaboración y aplicación de exámenes ha respondido más a fines de selección, que de enseñanza o aprendizaje

De igual manera, en la "era de la examinación a gran escala" podemos identificar ciertas ausencias tales como

- Necesidad de no sólo medir, sino evaluar y retroalimentar la acción educativa en las escuelas
- Necesidad de conocer la situación no sólo del producto, sino del proceso educativo
- Necesidad de una retroalimentación en materia de didáctica a partir de la medición realizada

Todo esto permite vislumbrar la transformación del sentido que ha sufrido el examen, sobre todo en su modalidad a gran escala y el impacto social, más que pedagógico que éstos han tenido tanto en el ámbito nacional como internacional, lo que revela la necesidad de darle un sentido más pedagógico al uso de los mismos

#### **1.4 Impacto sociopolítico y transformación del sentido de los exámenes a gran escala.**

La transformación histórica que ha sufrido el examen en sus diferentes modalidades y aplicaciones, permite ubicar a los exámenes a gran escala como el elemento ideal para juzgar a los sistemas educativos y sus actores. El sentido didáctico que Comenio les atribuyó al considerarlos parte del método de enseñanza, ha quedado soslayado.

La revisión y demostración de lo aprendido, intenciones con que el examen escolar ingresa al ámbito educativo y la idea de impulsar una reforma educativa a través de exámenes a gran escala ha quedado atrás. Lo único que prevalece hoy es la obsesión para la aplicación de exámenes como una forma de controlar y diferenciar.

De ahí que lo más importante hoy en los sistemas educativos, sea pasar el examen. Ante lo cual, la tarea del maestro, si quiere resguardar su nombre y espacio, ya no es enseñar, sino preparar para el examen.

De hecho existen publicaciones difundidas por casas editoriales<sup>36</sup> e instituciones que por una mínima cantidad ofrecen las "Guías para el examen CENEVAL"<sup>37</sup>. Esto denota una pérdida del sentido educativo y formativo, ya no sólo del examen sino de la función docente y de los sistemas educativos en general.

Más que un impacto educativo, el examen, sobre todo en su modalidad a gran escala, ha tenido un impacto ya no sólo social, sino político y ahora hasta económico. La comercialización adquirida por los exámenes a través de las "guías elaboradas para tal efecto" y la corrupción generada por el uso o mal uso que se le ha dado al examen, sobre todo en su modalidad a gran escala, ha ocasionado que el examen sea visto como el elemento idóneo, ya no para conocer, sino para enjuiciar la acción educativa de docentes y escuelas.

---

<sup>36</sup> Véase por ejemplo las guías para el examen ENLACE que ha publicado Santillana. El contenido de esta Guía la podemos encontrar en <http://www.santillana.com.mx/libro.php?ean=9789702920588>.

<sup>37</sup> El sitio web donde se pueden adquirir las Guías de CENEVAL resueltas es <http://ceneval.comerciomexico.com/>.

Consideramos que no es a través de exámenes que se va a lograr la tan ansiada calidad y mejora educativa en materia de enseñanza-aprendizaje, sino de la acción participativa de los actores que tienen a su cargo esta tarea: los docentes, sujetos que no sólo han sido olvidados en esta era de la examinación, sino enjuiciados y deshonrados.

La obsesión por la medición y el rendimiento de los alumnos en los sistemas, lleva a cuestionarnos sobre: ¿es éste el sentido de la escuela hoy?, ¿qué le dice al docente?, ¿cuál es el tipo de información que el docente requiere en materia de exámenes a gran escala para que en verdad éstos puedan decirle algo en relación a su práctica docente y aprendizaje de sus alumnos?

## Capítulo 2

### LOS EXÁMENES A GRAN ESCALA DESDE UNA PERSPECTIVA DIDÁCTICA

*Los exámenes sólo pueden aceptarse  
en cuanto comprueban la aptitud del niño para la vida social (J. Dewey)*

La aproximación al estado del arte en el que se encuentran las pruebas a gran escala y las corrientes de pensamiento desde las cuales se les ha estudiado, tema desarrollado en el capítulo anterior, permite identificar el predominio de enfoques centrados en la medición, la rendición de cuentas y la diferenciación social, para el estudio de pruebas a gran escala

El dominio de una perspectiva meramente cuantitativa, con visión administrativa y fines clasificatorios, que hasta la fecha han tenido los exámenes a gran escala, los coloca como instrumentos de medición, lejanos del aprendizaje y del sentido formativo que debe perseguir la evaluación del mismo

El objetivo del presente capítulo, es ofrecer una perspectiva alternativa para el estudio de las pruebas a gran escala y proponer un modelo de análisis, que desde una perspectiva didáctica, permita un manejo diferente para la evaluación a través de pruebas a gran escala. Al considerar los elementos que intervienen en el proceso de aprendizaje del alumno y su vinculación con situaciones de enseñanza, se pretende contar con los elementos necesarios que permitan darle un abordaje distinto a las pruebas a gran escala

Iniciamos con la justificación de la necesidad de una perspectiva didáctica para las pruebas a gran escala, al presentar las reacciones y perversiones suscitadas, las discusiones que desde la didáctica se dan y las experiencias existentes al respecto. Esto constituirá la base para la construcción de una perspectiva didáctica para los exámenes a gran escala

#### **2.1 Una mirada didáctica para los exámenes a gran escala**

Una mirada didáctica para los exámenes a gran escala, sólo podrá generarse dependiendo de la función que éstos cumplan de su vinculación con las actividades del aula para generar procesos de enseñanza y aprendizaje y del uso que se les dé a los resultados obtenidos



El no considerar al examen como parte del método didáctico y la ausencia de un manejo más formativo de estos instrumentos de evaluación, ha ocasionado una serie de reacciones y vicios en torno a los exámenes a gran escala, por el uso o más bien mal uso que a éstos se les ha venido dando

### **2.1.1. Reacciones y debates en el uso de los exámenes a gran escala como Instrumento de evaluación**

Las reacciones más fuertes en torno al examen como instrumento de evaluación, son de diverso tipo. Algunas están enfocadas a las medidas perversas que se toman en reacción al examen como instrumento de evaluación y otras se refieren a los problemas ocasionados en la interpretación de resultados, por el desconocimiento de lo que significa un examen estandarizado, sus posibles usos e implicaciones.

Existen estudios y medidas que demuestran que la validez de las pruebas estandarizadas para medir el logro del estudiante se puede ver comprometida, cuando los profesores “enseñan para la prueba”, por el temor a las consecuencias que les puedan acarrear, los resultados obtenidos (Sanders y Horn, 1995). Cuando los resultados de las pruebas se convierten en la meta de la enseñanza, el acto educativo se arruina. Se reporta el caso de un profesor de geometría en Nueva York que había sido reconocido con base al resultado que sus estudiantes obtuvieron en el examen nacional de geometría. Sin embargo, desgraciadamente, el funcionamiento de sus estudiantes se debió a la memorización que éstos hicieron de las doce pruebas que incluía el examen nacional.

Aquí en México, no es la excepción, las medidas adoptadas en torno a diversos exámenes a gran escala de circulación nacional, distorsionan por completo el sentido de un examen. Algunas de estas acciones, como la venta de guías resueltas para los exámenes de egreso del CENEVAL, que permitirán obtener el resultado deseado sólo con memorizarlas,<sup>38</sup> o bien la publicación por parte de editorial Santillana de manuales para contestar el examen ENLACE,

---

<sup>38</sup> Para mayor información, se puede consultar los siguientes sitios web: <http://www.guiasdeceneval.com/> y [http://ceneval.comercio.com/exco.com/?qclid=CPHxwBLE9psCFQJvswodmDQ8\\_w](http://ceneval.comercio.com/exco.com/?qclid=CPHxwBLE9psCFQJvswodmDQ8_w)

denominados "Aprobado" funcionan como una preparación para aprobar el examen ENLACE como su nombre lo indica.<sup>39</sup>

Por su parte, Casas (2003), expone la contradicción existente entre las exigencias que se le hacen al profesor con relación a la promoción de una enseñanza "auténtica",<sup>40</sup> o más bien centrada en la práctica, pero paradójicamente también le piden la aplicación de pruebas estandarizadas para evaluar el aprendizaje de sus alumnos.<sup>41</sup> La autora expresa que a pesar de que estandarización y comprobación del aprendizaje se han convertido en los vocablos por excelencia para valorar el trabajo de alumnos y docentes, ésta se vuelve una valoración forzada de contenidos que quizá no hayan sido asimilados aún, porque sólo miden la retención de conceptos y no evalúa el aprendizaje natural, además de causar gran ansiedad en los alumnos, al ser un requisito indispensable para obtener la acreditación y acceder al siguiente nivel de escolarización. Los estudiantes saben que su vida y éxito escolar se verá afectada por su desempeño frente a una prueba

Nichols y Berliner (2007), investigan las formas en que las pruebas ponen en peligro los fines e ideales del sistema educativo americano. Su análisis proporciona un marco global y una amplia gama de argumentos en contra de lo que ellos denominan "el gran juego de las pruebas" y los daños colaterales que éstas provocan, por el mal uso que de éstas se hace

Los mayores problemas en torno a la estandarización de la evaluación a través de pruebas a gran escala, tienen que ver con los sistemas de incentivos que recompensan o sancionan a las escuelas, según los resultados obtenidos por sus estudiantes, en los exámenes (Darling, 2003)

---

<sup>39</sup> Estas guías se pueden consultar en: <http://www.santillana.com.mx/libro.php?ean=9789702920588>

<sup>40</sup> Término utilizado por la corriente sociocultural del constructivismo social, para referirse a un aprendizaje situado y experiencial. Cabe destacar que los antecedentes y sustentos de este aprendizaje centrado en la práctica y en experiencias reales, lo podemos encontrar en la obra e ideas de John Dewey. En el marco del movimiento escuela nueva que surge como parte de la educación progresiva en Estados Unidos. Dewey (1958), presenta como una de sus tesis básicas la vinculación entre experiencia y educación, fundada en una filosofía empirista que lleve a una educación progresiva. La experiencia para Dewey es el medio principal de la educación, pero no se trata de cualquier tipo de experiencia, sino de experiencias que cumplan con las cualidades básicas señaladas por John Dewey, lo que constituye en sí la teoría de la experiencia propuesta por él (Dewey, 1958). Para ampliar la información sobre los antecedentes del enfoque experiencial y las prácticas educativas auténticas desde una perspectiva situada (véase Díaz Barriga Arceo, 2005).

<sup>41</sup> En México no es la excepción, pues mientras a partir del ciclo escolar 2009-2010, se impulsa una Reforma por competencias, el aprovechamiento y trabajo de los docentes se evalúa por pruebas a gran escala, centradas en la acumulación y reproducción de información,

En México, ésta es una medida comprendida en la llamada *Alianza Nacional para la Calidad de la Educación*,<sup>42</sup> establecida por la Secretaría de Educación Pública en mayo del 2008<sup>43</sup> en la cual se establecen bonos de estímulos para los maestros cuyos alumnos superen el promedio nacional en la prueba nacional censal aplicada desde el 2005 denominada *Evaluación Nacional de Logro Académico en Centros Escolares* mejor conocida como (ENLACE), que precisamente constituye el objeto de estudio de esta investigación

De acuerdo con algunos investigadores y analistas al respecto, entre los efectos perversos de esta medida, se encuentra que los maestros ya no se preocupan por enseñar, sino por entrenar a los alumnos para la prueba. Dice Díaz Barriga (2008), si los maestros son recompensados por los resultados que obtengan sus alumnos en la prueba ENLACE, el proyecto pedagógico de este sexenio quedará definido por dicha prueba. Esto es, en vez de contenidos, de programas escolares o de formación en competencias, los maestros se dedicarán a entrenar a sus alumnos en el llenado de bolitas (*la prueba de bolitas le dicen los niños en la escuela*), porque en ese resultado se juega su estímulo

Esta medida no toma en cuenta, estudios anteriores realizados en la Unión Americana entre 1940 y 1950, que demostraban que la estrategia de pagar a los profesores de acuerdo a los resultados de los estudiantes traía consigo muchos problemas para su aplicación

Lo más grave, es la descontextualización de resultados en las medidas que se toman y la omisión del contexto socioeconómico y cultural en que se dan estos resultados

Para Muñoz Izquierdo (2008)

“al distribuir los estímulos al desempeño de los docentes con base en el aumento que éstos logren, de un año a otro, en los puntajes que sus estudiantes obtengan en la prueba *ENLACE*, será indispensable estratificar socio-académicamente a los alumnos, para que todos los maestros compitan en las mismas circunstancias. Si esto no se

---

<sup>42</sup> El programa establece crear estímulos a la calidad docente, estimulando el mérito individual de los maestros en función exclusiva de los resultados del logro de sus alumnos (SEP, 2008a).

<sup>43</sup> Para mayor información se puede consultar SEP (2008a) “Alianza Por la calidad de la educación” Boletín No 119 16 de mayo 2008 Consultado en: <http://www.sep.gob.mx/web/sep1/bol1190508>

---

hiciera, sólo se premiaría a quienes trabajen en los ambientes más favorables para el aprendizaje, en detrimento de los docentes que desempeñen su profesión en ambientes socioeconómica y culturalmente adversos” (p. 1)

Por su parte Díaz Barriga (2008) alude a que no se puede juzgar de igual manera a un maestro que trabaja en condiciones adversas que a quienes cuentan con recursos humanos, materiales y culturales más propios. Un maestro que trabaja con grupos vulnerables y en situaciones de reprobación, difícilmente podrá obtener los mismos resultados que quien cuenta con todos los recursos y cuyo trabajo se desarrolla en situaciones más favorables. Lo que no quiere decir que el trabajo de este maestro sea menor.

Las dificultades ocasionadas por un mal manejo de los resultados en un test, han sido estudiadas por autores como Ferrer y Arregui (2003), Tiana (1997) y Ravela (2006). Entre éstas se pueden citar, el ranqueo en la presentación de resultados, el desconocimiento de cómo y desde dónde interpretar los resultados obtenidos en una prueba y la ignorancia de la desigualdad de condiciones socioculturales y económicas existentes en los distintos contextos en donde se aplica una prueba a gran escala. Es decir de la dificultad de homogenizar el aprendizaje y sus resultados ante una heterogeneidad de condiciones en donde se realiza la enseñanza.

El desconocimiento de ¿qué son? y ¿para qué sirven las evaluaciones educativas?, de los elementos básicos de un proceso de evaluación, de la conceptualización y enfoque del objeto a evaluar, de lo que evalúa una prueba de acuerdo a los distintos tipos de actividades, de lo que significan los números en las evaluaciones, de los problemas comunes a todas las evaluaciones, de los diversos significados que se le pueden dar a la rendición de cuentas en sus distintas versiones y manejos, como es la visión errada e incompleta que puede ocasionar un ranqueo educativo. Así como la ignorancia de los factores asociados<sup>44</sup> a todas las evaluaciones. Además del mal uso que a éstas se les da al privilegiar el reporte estadístico en los informes de evaluación, y la omisión de informes con información valiosa y útil para los destinatarios, más centrada en el aprendizaje y la enseñanza que en la medición.

---

<sup>44</sup> En América Latina se suele utilizar la denominación “estudios de factores asociados” a los trabajos de análisis e investigación desarrollados a partir de la interrelación entre resultados de pruebas e información recogida principalmente a través de los cuestionarios complementarios (Ravela, 2006k).

En el Cuadro 7 se presenta una síntesis de los problemas más apremiantes, que de acuerdo con diversos autores, se identifican tanto en el ámbito estadounidense como en el latinoamericano y nacional en cuanto al uso de los tests como instrumento de evaluación.

Cuadro 7

<b>Reacciones y debates en el uso del test como instrumento de evaluación.</b>		
<b>Tipo de problemas</b>	<b>Autores</b>	<b>Problemas detectados</b>
Medidas	Sanders y Horn (1995)	Los profesores enseñan para la prueba
	Casas (2003)	Contradicción entre promoción de nuevas formas de enseñanza y comprobación forzada de contenidos a través de pruebas. En donde sólo se mide la retención de conceptos y no el aprendizaje natural
	Darling (2003)	Sistema de incentivos que recompensan o sancionan a las escuelas
	Nichols y Berling (2007)	Daños colaterales
	Díaz Barriga (2008)	Entrenar a los alumnos para contestar la prueba en vez de enseñar
	Muñoz Izquierdo (2008)	Asignación de estímulos por los resultados obtenidos, ignorando el contexto socioeconómico y cultural en que éstos se obtienen
	Tiana (1997)	Desconocimiento de cómo y desde dónde interpretar los resultados
	Ferrer y Arregui (2003)	Rankeo en la presentación de resultados  Ignorancia de la desigualdad de las condiciones culturales en cada país

Interpretación y uso de resultados	Ravela (2006b)	Desconocimiento de ¿qué son? Y ¿para qué sirven las evaluaciones educativas?, de los elementos básicos de un proceso de evaluación, de la conceptualización y enfoque del objeto a evaluar, de lo que evalúa una prueba de acuerdo a los distintos tipos de actividades, de lo que significan los números en las evaluaciones, de los problemas comunes a todas las evaluaciones, de los diversos significados que se le pueden dar a la rendición de cuentas en sus distintas versiones y manejos y de los factores asociados
------------------------------------	----------------	--

De acuerdo con lo que se muestra en el cuadro 7, podemos identificar que los principales problemas en torno a las pruebas a gran escala como instrumento de evaluación, residen en los efectos perversos que éstas producen, sobre todo por el mal manejo que de éstas se hace y un desconocimiento de lo que significan los resultados, así como por la visión errónea de la realidad educativa que presentan los rankeos que de las pruebas se hacen. Sobre todo, porque los destinatarios y agentes encargados de elaborarlos y difundirlos, son sujetos ajenos al proceso educativo y de enseñanza.

Las reacciones y debates en torno al uso de los tests y los problemas que éstos ocasionan, dan cuenta de la necesidad de abordar al examen desde un enfoque más didáctico; es decir, no sólo desde sus problemas, sino también desde sus posibilidades como generadores de actividades de enseñanza en el aula, e instrumento del docente para corregir el error.

### **2.1.2. El examen como instrumento de evaluación desde un enfoque didáctico**

Un principio didáctico fundamental es la estrecha relación que debe existir entre formas de enseñanza, procesos de aprendizaje y sistemas de evaluación. Esto requiere que la manera de enseñar, esté en estrecha relación con la manera de evaluar, así como con el objeto de evaluación y el propósito de la misma. Desde una perspectiva didáctica, no es posible evaluar lo que no se enseña.

De igual manera, si lo que quiero es evaluar el aprendizaje, es importante primero definir ¿qué entiendo por aprendizaje? o ¿qué tipo de aprendizaje quiero evaluar? y ¿desde qué postura o concepción de aprendizaje se va a evaluar, aplicar y usar el instrumento de evaluación? Es decir, si se concibe el aprendizaje y su evaluación, desde una postura conductista, donde lo más importantes es el almacenamiento y reproducción mecánica de información. O bien, de acuerdo con las nuevas tendencias como un modo de proceder reflexivo, vinculado a la vida misma (Kilpatrick, 1968), como producción de significados que se da por la interacción entre identificación y utilización (Meirieu, 1992), como reinterpretación de significados y construcción de una cultura escolar mediante procesos de intercambio y negociación (Pérez Gómez, 2000), o como movilización de saberes (Perrenoud, 2002).

Otra cuestión, es pensar en las características y enfoques de las situaciones de enseñanza que se dan en el aula. Es decir, en el tipo de actividades de aula, mediante las cuales se promueve el aprendizaje. Éstas deberán estar en estrecha relación con el enfoque de aprendizaje y ser la base para posteriormente construir las situaciones o actividades de evaluación. Por ejemplo, si se piensa en la enseñanza como mera transmisión de información, tendremos actividades de aula centradas en la recirculación de la información y la evaluación tendrá que ir encaminada a la reproducción de la misma. Si por el contrario, de acuerdo con lo reportado por Meirieu (1992), consideramos a la enseñanza como descubrimiento de los puntos de apoyo en el estudiante, las actividades del aula, tendrán que considerar estos puntos de apoyo, de tal manera que sean la base para que el estudiante encuentre sentido y significado a lo que está aprendiendo a través de la movilización de sus saberes. Esto implica que las actividades de evaluación, deberán tener también un significado para el alumno y constituir una oportunidad para demostrar lo aprendido mediante la resolución de problemas.

De acuerdo con el estudio realizado por Bain (2006) la evaluación debe ser un medio más para ayudar a aprender, ante lo cual los exámenes deberían ser parte importante del trabajo del curso y no meros instrumentos para clasificar y jerarquizar.

El Cuadro 8 muestra la función de los exámenes como instrumento de evaluación, desde distintas perspectivas de aprendizaje

Cuadro 8

<b>El examen como instrumento de evaluación desde diferentes concepciones de aprendizaje</b>				
<b>Perspectiva</b>	<b>Aprendizaje</b>	<b>Enseñanza</b>	<b>Evaluación</b>	<b>Examen</b>
Conductista	Almacenamiento de información	Conexión de estímulos	Reproducción de la información	Instrumento de verificación
Significativa	Producción y reinterpretación de significados	Conexión entre identificación y utilización Procesos de intercambio Descubrimiento de puntos de apoyo	Oportunidad para demostrar lo aprendido, movilizandosaberes mediante la resolución de problemas	Medio para ayudar a aprender a través de la corrección del error

Esto lleva a preguntarnos, en un sistema de evaluación a través de exámenes ¿Cuál es el concepto de aprendizaje y enseñanza que se promueve? ¿Qué tanto ayuda a la movilización de conocimientos y la aplicación del mismo, para ver si los alumnos lograron un aprendizaje más significativo? ¿Qué tanto le ayudan al docente a descubrir los puntos de apoyo del estudiante y los elementos que deberá reforzar?

La reflexión generada de estos cuestionamientos, puede ser la base para descubrir alcances y limitaciones que desde un enfoque didáctico pueden presentar los exámenes a gran escala como instrumentos de evaluación

### **2.1.3 Alcances y limitaciones de las pruebas a gran escala como instrumentos de evaluación**

Muchos autores, desde la implementación del examen como instrumento de evaluación de la institución escolar (exámenes escolares) o de un sistema educativo (exámenes a



gran escala) han cuestionado, no el uso del examen sino el mal uso que a éstos se les está dando, al empearlos como instrumentos clasificatorios que no contribuyen a la mejora de la educación y sobre todo de la enseñanza. Siendo ésta una de sus principales limitaciones.

Los integrantes del Comité de la Sociedad Nacional para el Estudio de la Educación, ya condenaban desde mediados del siglo XX, la evaluación de los productos del esfuerzo educativo a través de exámenes administrativos que sobrevaloran la memoria de hechos y principios y descuidaban los resultados más dinámicos de la instrucción (Kilpatrick, et al. 1944). Dice Kilpatrick (1968) "Nuestra preocupación no debe ser que el niño aprenda las materias de suerte que pueda rendir exámenes sobre ellos" (p. 51).

Otros estudios realizados (Wildemuth, 1984) revelan que la mayoría de las veces las pruebas a gran escala, no necesariamente prueban lo que han aprendido los estudiantes. Además de que no constituyen un elemento central para mejorar la acción docente en el aula.

Cimbricz (2002), examina la relación existente entre pruebas nacionales (llamadas pruebas de Estado) aplicadas de manera obligatoria por el Estado y las creencias y prácticas de los profesores. Los estudios presentados en este artículo revelan que las pruebas aplicadas por el Estado, ejercen importante influencia en lo que los profesores dicen y hacen, pero no en cuanto a la enseñanza. La influencia que las pruebas nacionales ejercen sobre el profesor depende de cómo los profesores interpretan estas pruebas, por lo tanto la prueba no aparece como un elemento que impulse el cambio, porque no se presenta y trabaja desde esta perspectiva.

Tenemos que el problema no reside tanto en utilizar el examen como instrumento de evaluación, sino en el manejo que de éstos se hace. Un manejo más pedagógico del examen con enfoque formativo, centrado no tanto en resultados, sino en procesos implícitos, que más que cuantificar el logro alcanzado ayude a mejorar los procesos de enseñanza-aprendizaje y la acción docente dentro del aula, ayudara a identificar y aprovechar algunas de las posibilidades que este instrumento de evaluación puede contener.

De acuerdo con algunos autores (Baudillo, 2008) la evaluación del aprendizaje, el rendimiento y el desarrollo de capacidades, deberán constituir una oportunidad para profundizar sobre qué, como y para qué aprende el educando. Es una reflexión pedagógica válida que debe llevar a la confrontación, interlocución y actualización didáctica.

Creemos que un manejo más didáctico del examen a gran escala tanto en su elaboración, como en su operación y uso, podría ayudar a estructurar y entender el proceso evaluativo a través de exámenes a gran escala, desde una perspectiva diferente.

Es decir, la evaluación del aprendizaje a través de exámenes a gran escala con un enfoque educativo y didáctico, debe ser concebida y desarrollada como un acto de análisis y reflexión para la acción docente, que permita obtener información de cómo está aprendiendo el estudiante, qué es lo que estás aprendiendo, en qué nivel de competencia se ubica, que tipo de conocimientos ha desarrollado y que tipo de medidas correctivas o estrategias se pueden implementar para estimular a los estudiantes en riesgo a superar las dificultades.

Así, tenemos que entre los principales alcances y limitaciones que pueden tener los exámenes a gran escala desde una perspectiva didáctica, se encuentran los siguientes:

Cuadro 9

**Los exámenes a gran escala desde una perspectiva didáctica.  
Alcances y limitaciones**

Alcances	Limitaciones
Acto de análisis y reflexión para la acción docente  Permite obtener información de cómo está aprendiendo el estudiante, qué es lo que está aprendiendo, en qué nivel de competencia se ubica y que tipo de conocimientos ha desarrollado  Ayuda a implementar medidas correctivas o estrategias se pueden implementar para estimular a los estudiantes en riesgo a superar las dificultades	Instrumento clasificatorio  Privilegian la memoria de hechos y principios  No necesariamente reflejan lo que los estudiantes han aprendido  No aparecen como un elemento que impulse el cambio  No constituyen un elemento central para mejorar la acción docente en el aula

La superación de las limitaciones presentadas en el Cuadro 9 y la obtención de los alcances aquí enunciados, sólo podrá verse cuando los exámenes sean utilizados como elementos que permiten la reflexión del docente sobre su práctica del aula, a través de análisis de las actividades de evaluación y de los errores que los alumnos cometen en la respuestas emitidas, es decir cuando se le de al examen a gran escala un manejo más didáctico

#### 2.1.4 Estudios y experiencias sobre un manejo didáctico de los exámenes a gran escala

Para tener una idea del manejo didáctico que puede dársele tanto a los exámenes a gran escala como a los exámenes del aula, podemos recurrir a dos experiencias en el ámbito Latinoamericano la de la Administración Nacional de Educación Pública (ANEP) en Uruguay a cargo de la Dirección de Investigación, Evaluación y Estadística, (DIE) dependencia encargada de coordinar el sistema de evaluación<sup>45</sup> en Uruguay a través de pruebas nacionales<sup>46</sup> e

<sup>45</sup> Sistema que se estructura a partir de 1996, cuyo principal objetivo es usar la evaluación con fines formativos para una mejora de la práctica docente. En donde lo que interesa no es cuantificar el logro alcanzado, sino mejorar el sistema. El objetivo principal es producir información de calidad con rigor metodológico que sirva para la toma de decisiones

<sup>46</sup> Conviene resaltar que para el diseño de estas pruebas se considera el objetivo final y parcial de la misma, el grado o nivel, el momento de aplicación, los aspectos de la asignatura a evaluar y el marco teórico de la actividad

internacionales,<sup>47</sup> y la de un trabajo argentino con docentes, en donde la evaluación a través de exámenes a gran escala adquiere un sentido más didáctico

La experiencia uruguaya a cargo de la ANEP reporta tres tipos de actividades, que tienen por objeto darle un manejo didáctico a la evaluación del aprendizaje a través de pruebas a gran escala, tanto a nivel nacional como internacional

La primera de éstas reside en diversas publicaciones tales como Informes sobre los resultados de evaluaciones nacionales<sup>48</sup> e internacionales como PISA.<sup>49</sup> Publicaciones sobre las áreas evaluadas, las cuales se caracterizan por presentar además de las actividades de prueba y los resultados obtenidos, el marco de evaluación que las sustenta. Boletines de difusión electrónica que analizan los contenidos y procesos cognitivos implícitos en las evaluaciones de PISA y algunos estudios sobre clima educativo. También cuenta con un sitio web en donde podemos encontrar además de las publicaciones mencionadas, actividades de prueba y bases de datos de los estudios sobre PISA.<sup>50</sup>

---

Por ejemplo la "Evaluación nacional de aprendizajes en lenguaje y matemáticas para 6° año de enseñanza primaria 2005", va acompañada de un documento que explica el enfoque que se seguirá en dicha evolución, el diseño y descripción de las pruebas, las características sociales de la población evaluada y una guía para interpretar los resultados (ANEP, 2007c). Cada una de estas evaluaciones cuenta una etapa previa de planificación, organización, comunicación a los actores y juicio. Una de ejecución para la recolección de datos, una de procesamiento y análisis de datos y una última de uso de la información.

<sup>47</sup> Entre éstas, están el *Segundo Estudio Regional Comparativo y Explicativo (Valdés, 2008)*. Estudio aplicado en 2006 por el Laboratorio Latinoamericano de Evaluación de la Calidad de la Educación (LLECE) de la Oficina Regional de Educación de la UNESCO para América Latina y el Caribe. Tiene el propósito generar conocimiento sobre los rendimientos de los estudiantes de 3ro y 6to año de Educación Primaria en América Latina y el Caribe en las áreas de Lengua (lectura y escritura), Matemática y Ciencias e identificar los factores que tienen influencia y/o explican los desempeños de los estudiantes en estas áreas. Y la prueba PISA (Programa Internacional para la Evaluación de Estudiantes (en inglés, Programme for International Student Assessment)). Este programa tiene como principal objetivo evaluar en qué medida los sistemas educativos de los países participantes han logrado preparar a los estudiantes que están terminando el período de educación obligatoria para la participación plena en la sociedad del conocimiento. La aplicación, procesamiento y difusión de la información de este estudio en Uruguay es coordinado por el programa *PISA en Uruguay* de la ANEP.

<sup>48</sup> Por ejemplo el informe de la "Evaluación piloto autónoma para primer año de educación primaria en Lectura y escritura" cuya característica más destacable incluye la descripción de los niveles de desempeño en lectura y escritura, ejemplos de pruebas de cada nivel y aportes a la enseñanza de la lectura y escritura. (ANEP-CÓDIGEN, 2008)

<sup>49</sup> El más reciente informe sobre PISA 2006 en Uruguay (ANEP, 2007), comprende además del marco conceptual de PISA y la descripción de la participación de PISA en Uruguay, un mapa de la desigualdad educativa en Uruguay, con el fin de contextualizar los resultados y un estudio de caso sobre las características de los sistemas educativos y de formación de docentes de los países o regiones que durante tres años consecutivos han obtenido desempeños destacables en PISA (Finlandia, Korea y Hong Kong) en las áreas de lectura, matemáticas y ciencias. Esto con el propósito de conocer cuáles son los factores que inciden en los resultados y tener la posibilidad de aprender de ellos. Cabe destacar además, la existencia de publicaciones donde se presenta y analiza el marco de evaluación de PISA a partir de algunas actividades de prueba como *Resolución de problemas para el mundo de mañana. PISA 2003. Marco Conceptual y Actividades de prueba* (ANEP, 2006) y *La evaluación de la "Cultura científica" en PISA 2003. Marco Conceptual y Actividades de prueba* (ANEP, 2004)

<sup>50</sup> <http://www.anep.edu.uy/sitio/anep.php?identificador=212>

Otra de las actividades consiste en la creación de Comités Técnicos, con representantes de todos los niveles educativos, especialmente docentes. Éstos constituyen un grupo de especialistas que generan reflexión y producen conocimiento con el objetivo de estudiar, discutir y difundir en sus ambientes el marco de la evaluación PISA y cuyo propósito es generar materiales de trabajo como aporte a los docentes a partir de las actividades de prueba.

El tercer tipo de actividad y que resulta ser de las más significativas, para los fines de esta investigación, son los talleres para docentes para el análisis de las actividades de PISA. Estos talleres, dirigidos a docentes, constituyen un marco de reflexión sobre las prácticas de aula desde el marco de evaluación de PISA. Entre las principales actividades se encuentran la lectura y discusión del marco conceptual de PISA. El análisis didáctico de las actividades realizadas en donde se le pide al docente que resuelva la actividad de prueba y después identifique en estas actividades los contenidos, los procesos implícitos y el contexto en el que se ubican. Y a partir de aquí diseñen actividades del aula que puedan ayudar a desarrollar los procesos identificados, a partir de lo cual los docentes se dan cuenta que los contenidos son el vehículo para el desarrollo de las competencias en los alumnos. Finalmente se les pide a los docentes diseñen actividades de prueba semejantes, cambiando y mejorando las preguntas revisadas.

Lo más interesante de la experiencia uruguaya, en el manejo didáctico de las pruebas a gran escala, es sin duda, la publicación de los marcos teóricos y conceptuales de los instrumentos generados, el análisis de los contenidos y procesos implícitos en las actividades de prueba y la generación y diseño de actividades del aula que potencialicen el desarrollo de esos procesos.

En Argentina existe un trabajo realizado por Bertoni, Poggi y Teobaldo (1996) que consiste en un proyecto de investigación evaluativa realizado con docentes de 42 escuelas primarias piloto en donde se reporta, como a partir de tres preguntas básicas en torno a ¿qué significan los resultados obtenidos en las pruebas?, ¿a qué se deben estos resultados?, ¿qué debemos hacer?, se pretende devolver la evaluación a los docentes, hacer una

---

---

especie de evaluación de la evaluación y analizar los vínculos entre prácticas de enseñanza y prácticas de evaluación

La metodología empleada fue de corte tanto cuantitativo como cualitativo. Con tres niveles de análisis: contexto, saberes enseñados, saberes aprendidos.

Las principales aportaciones de esta investigación residen en las reflexiones sobre el problema de la evaluación educativa y los dilemas teórico-metodológicos que suscitan sus diferentes enfoques, proponiendo caminos alternativos y la posibilidad de optar por distintos modelos o enfoques.

También cabe destacar la participación de los diferentes grupos involucrados en el proyecto y las distintas tareas que cada uno realizó. En el caso de los investigadores, éstos cumplieron con el rol de facilitadores externos, estableciendo con los docentes relaciones de tipo cooperativo, para articular las reflexiones de éstos con la información recogida. Los supervisores, como gestores reales del proyecto participan activamente en su preparación y en la discusión de sus resultados. Los directores actúan como interlocutores críticos, principalmente en el análisis de procesos y resultados de la evaluación. Y los docentes tuvieron una importante participación en la elaboración de ítems y corrección de pruebas.

El trabajo realizado revela así las distintas implicaciones de la evaluación y la articulación que debe existir entre la teoría y la práctica, así como entre la tarea de control y el trabajo sobre las nuevas significaciones que los actores involucrados en el proceso construyen y reconstruyen en relación con ella. Lo importante era encontrar los referentes a partir de los cuales se realiza la tarea de la evaluación y la importancia que éstas adquieren tanto para la institución, como para el docente, como para el alumno. Considerando así que la evaluación desde esta perspectiva aborda tanto el plano cultural, como el teórico y práctico de los saberes enseñados y de los saberes aprendidos en la institución escolar.

Este trabajo pone de manifiesto los distintos significados que la evaluación del aprendizaje adquiere para los diferentes actores involucrados en ella. Y procura que a través de conocer lo que los supervisores, directores y docentes tienen que decir con respecto a ésta, que son los encargados de evaluar todos los días, de manera directa o indirecta. Todo esto con el fin de intentar que para los alumnos, la evaluación, no sea injusta ni piadosa, sino reveladora.

---

tanto de sus saberes, como de sus dificultades y posibilidades. Es decir la intención principal del trabajo realizado en esta investigación es abrir la posibilidad de que la evaluación constituya un acto educativo de profunda intencionalidad pedagógica y no un castigo ni una mala palabra (Bertoni, Poggi y Teobaldo, 1996)

Estamos ante dos trabajos de evaluación con enfoque formativo, en donde de acuerdo con Perrenoud (2008), la evaluación se pone más al servicio del aprendizaje que de la selección “No es una idea nueva que la evaluación pueda ayudar al aprendizaje del alumno. Pero no se puede mejorar la evaluación sin tocar el conjunto del sistema didáctico y del sistema escolar” (p. 8-9)

En el Cuadro 10 se presenta una síntesis de estas experiencias. Lo que permite visualizar de mejor manera el aporte de las mismas.

Cuadro 10

**Experiencias latinoamericanas de un manejo didáctico de los exámenes a gran escala**

País	Autor	Tipo de actividad	Propósito	Contenido
Uruguay	ANEP-DIE	Difusión	Elaboración de informes sobre resultados de evaluaciones nacionales e internacionales  Publicaciones sobre las áreas evaluadas	Marco de evaluación Actividades de prueba
		Talleres para docentes	Marco de reflexión sobre las prácticas de aula desde el marco de evaluación de PISA	Lectura y discusión del marco conceptual de PISA Análisis didáctico de las actividades realizadas Resolución de la actividad de prueba Identificación de las actividades, contenidos, procesos implícitos, y contexto en el que se ubican Diseño de actividades del aula que puedan ayudar a desarrollar los procesos identificados Diseño de actividades de prueba semejantes
Argentina	Bertoni, Poggio y Teobaldo	Investigación	Devolver la evaluación a los docentes Hacer una evaluación de la evaluación Analizar los vínculos entre prácticas de enseñanza y prácticas de evaluación	¿Qué significan los resultados obtenidos en las pruebas? ¿A qué se deben éstos resultados? ¿Qué debemos hacer?



Como vemos, el eje principal de estos estudios resulta ser el trabajo con los docentes para el análisis de las implicaciones didácticas y sentido que tienen las actividades de una prueba, lo cual constituye la base para la construcción de una perspectiva didáctica para el análisis de los exámenes a gran escala

### **2.1.5 Implicaciones en la construcción de una perspectiva didáctica para la evaluación del aprendizaje y los exámenes a gran escala**

La enseñanza, principal objeto de estudio de la didáctica, constituye un proceso intencional que incorpora elementos tanto de orden social, como curricular y personal

La didáctica como disciplina teórica, política e histórica, aporta elementos para una conceptualización y generación de una enseñanza facilitadora del aprendizaje que responda a las situaciones y necesidades tanto sociales como institucionales y personales en un momento histórico determinado

Así la didáctica, constituye una importante óptica para reflexionar sobre las implicaciones y acciones que en materia de enseñanza se pueden generar a partir de un marco social, curricular, institucional y personal

La evaluación, como parte fundamental de la acción didáctica, puede y debe ser pensada desde nuevas perspectivas y acciones. En donde enseñanza, aprendizaje y evaluación presenten una relación estrecha tanto en el orden conceptual como operativo, situacional y estratégico.

La evaluación del aprendizaje debe ser parte fundamental de la didáctica. Desde esta perspectiva, no constituye sólo una acción a realizar al finalizar el proceso de enseñanza-aprendizaje, sino que se convierte en parte esencial del mismo proceso. Y constituye un elemento fundamental para propiciar nuevos y mejores aprendizajes a través de la enseñanza y dentro de un contexto sociocultural determinado. Y en donde todos los actores involucrados y en especial el docente tienen una participación fundamental

Desde una perspectiva psicométrica los tests, están constituidos para demostrar el trabajo escolar, pero desde una mirada pedagógica y didáctica, los tests pueden ser un excelente medio para analizar e impulsar el trabajo escolar.<sup>51</sup>

Así, la evaluación deja de ser un proceso externo al aula, al docente y al estudiante, para convertirse en una acción que regresa al aula y forma parte de las actividades aquí realizadas. Porque proporciona información importante tanto al docente, como al alumno y a la institución sobre el impacto de la enseñanza y los ajustes y mejoras que a ésta se tienen que realizar a nivel de aula y curricular. Así mismo constituye un importante medio de conocer las habilidades cognitivas desarrolladas en los estudiantes y las que faltan por desarrollar.

Es importante considerar que para que un examen, ya sea a nivel del aula o a gran escala, pueda mirarse como un elemento potencializador del aprendizaje y de la enseñanza, es necesario que en su diseño y elaboración se incluyan no sólo contenidos, sino sobre todo las habilidades a evaluar. De tal manera que las actividades o tareas de evaluación movilicen saberes y no sólo soliciten la reproducción de información, sino la aplicación de la misma a situaciones reales.<sup>52</sup> Además de que se incluyan actividades con diferente grado de dificultad de tal manera que se puedan identificar los diferentes niveles de desempeño que pueden lograr los estudiantes.<sup>53</sup>

Aunque insistimos en que un examen difícilmente puede considerarse como lo más óptimo para evaluar ejecuciones en situaciones reales. Para algunos autores (Casas, 2003), el uso

---

<sup>51</sup> En países como Brasil e Inglaterra, se han desarrollado las llamadas *Matrices de referencia*, o niveles de desempeño a partir de las cuales, se analizan las habilidades necesarias para abordar los contenidos curriculares, lo que a su vez, puede servir como base para elaborar los tests y el trabajo escolar en el aula y con los alumnos (Fontanive, 2005).

<sup>52</sup> Como es el caso de los llamados exámenes de tercera generación (Díaz Barriga, 2007) como lo son los de PISA, cuya característica principal es la contextualización de los ítems en situaciones reales. En Padilla (2005) y Díaz Barriga, (2006) podemos encontrar estudios y análisis sobre las diferencias existentes en el contenido de los exámenes de corte internacional. El análisis comparativo del contenido y estructura de exámenes internacionales como PISA y su relación con exámenes nacionales como el *Examen Nacional de Nuevo ingreso a Educación Media (EXANI-1)* del Centro Nacional de Evaluación (CENEVAL), ayuda a identificar el nuevo concepto de examen que ofrecen los estudios de evaluación internacional a gran escala como PISA y los retos existentes en México para la elaboración de exámenes.

<sup>53</sup> Conviene resaltar que los exámenes elaborados por el equipo técnico-pedagógico de la ANEP y la DIE en Uruguay, ayudan a identificar después de su aplicación diferentes niveles de desempeño. Véase por ejemplo "Los niveles de desempeño al inicio de la educación primaria. Estudio de las competencias lingüísticas y matemáticas" (ANEP, 2002b).

de exámenes estandarizados para evaluar aprendizajes llamados “auténticos” o con la inclusión de situaciones reales desde una perspectiva más didáctica es una contradicción

De ahí la importancia de no privilegiar al examen normativo, sobre todo en su modalidad a gran escala, como único instrumento de evaluación y considerar los diversos enfoques y opciones que existen para la evaluación de un aprendizaje ubicado en situaciones reales. Lo cual puede resultar altamente significativo para los alumnos, porque a través de éstos, se pueda revisar lo aprendido, lo logrado y lo que falta por realizar a través de las llamadas rúbricas o matrices de valoración y de los portafolios.<sup>54</sup> Estrategias de evaluación que pueden resultar altamente útiles para lo que actualmente se ha dado en llamar el enfoque de competencias en la educación.

En el caso de las rúbricas o matrices de valoración, se trata de tablas de cotejo que describen los diferentes niveles en que se puede desarrollar una tarea. Estas pueden ser *holísticas o analíticas*. Las matrices holísticas describen procesos unidimensionales de los cuales se quiere tener una idea general, sin embargo brindan poca oportunidad de retroalimentación. Las matrices analíticas compuestas por tres elementos básicos: aspectos o dimensiones a valorar, niveles de desempeño y descripciones, tienen un propósito de evaluación más formativo, porque brindan mayor información que promueve la retroalimentación. Lo que permite a su vez crear un perfil de fortalezas y debilidades que se manifiestan en el desempeño de la actividad propuesta, a partir de lo cual estudiantes y docentes pueden elaborar el plan de mejora (Leymoní, 2008). Esto permite potencializar el aprendizaje y superación de los estudiantes.

En cuanto a los llamados portafolios, nombre que se le da a la elaboración de carpetas en la que se conjuntan diversas evidencias de trabajos realizados. Éstos pueden y deben ser analizados en función de los contenidos y habilidades desarrolladas en cada tarea. Cabe mencionar que si sólo se queda en la recopilación de trabajos, esto tiene mucho que ver con un principio didáctico fundamental de que la evaluación deberá recopilar varias evidencias de trabajo, y no tendrá nada de innovador (Díaz Barriga, 2009). Sin embargo, consideramos que su máxima potencialidad reside en que no sólo se trata de que la evaluación del aprendizaje se realice a través de los distintos trabajos realizados, sino que éstos deberán

---

<sup>54</sup> Nombre que desde el enfoque de evaluación auténtica (Díaz Barriga Arceo, 2005) se le da a lo que en didáctica se conoce con el nombre de “múltiples evidencias” (Díaz Barriga, 2009).

mostrar el avance logrado en materia de conocimientos y habilidades, o bien de competencias. Es decir, el conjugar los trabajos realizados en un curso y armar una carpeta con éstos, permite al estudiante darse cuenta de lo aprendido y trabajado durante el curso y analizar lo que aprendió con cada uno. Es decir cómo inició y cómo terminó. Por lo que un buen uso de los llamados portafolios como instrumento de evaluación permite más que la mera recolección de las evidencias de aprendizaje, la reflexión y análisis sobre lo aprendido en cada una de éstas.

Con respecto al enfoque de evaluación por competencias, a pesar de las múltiples definiciones que existen con respecto a las competencias, tres son los elementos que se pueden distinguir para integrar una competencia: conocimientos, habilidades y situaciones inéditas. De acuerdo con Perrenoud (2002) una competencia se define como "la capacidad de actuar de manera eficaz en un tipo definido de situación, capacidad que se apoya en conocimientos pero no se reduce a ellos" (p.7). Una evaluación por competencias implicaría entonces pensar en situaciones inéditas de la vida cotidiana, es decir en la resolución de problemas reales que requieran de la movilización de conocimientos y habilidades.

Estas alternativas constituyen una reacción a la evaluación centrada en resultados, que privilegia la medición de conocimientos declarativos. Considerando que para algunos autores (Ariasian 2001, Darling-Hammond, 1995, Díaz Barriga Arceo y Hernández, 2002, Díaz Barriga, 2005), la prerrequisita central de una evaluación a la que han llamado "auténtica" es que hay que evaluar desempeños y aprendizajes contextualizados. Una característica común en todas estas propuestas de evaluación es que sus desarrollos parten de la necesidad de evaluar desempeños y no sólo contenidos factuales, además de que está centrada en situaciones reales.

Tenemos entonces que sólo mediante el empleo y la combinación de diferentes alternativas de evaluación, se podrá lograr la tan ansiada mejora educativa; que en sí desde un enfoque formativo, vendría a constituir la función didáctica de la evaluación.

De ahí la necesidad de que tanto los exámenes sea a nivel de aula, o a gran escala, así como los distintos instrumentos de evaluación que se pueden utilizar, sean utilizados para retroalimentar la acción docente en materia de enseñanza y los avances y dificultades.

presentadas por los alumnos en su aprendizaje, pero siempre dentro de un contexto social determinado

Es importante tomar en cuenta que independientemente del instrumento utilizado para la evaluación del aprendizaje éste deberá contener tres elementos fundamentales

- Un marco conceptual que sustente y explique el contenido del objeto a evaluar, como de los propósitos de la actividad de evaluación
- Una actividad de evaluación que permita conocer y potencializar los procesos desarrollados y niveles de desempeño alcanzados
- Un modelo de interpretación que proporcione información sobre los logros y dificultades observadas en la actividad de evaluación

De esta manera la evaluación del aprendizaje ya sea a través de exámenes o de sus diferentes instrumentos, se convertirá en una estrategia de enseñanza empleada por el profesor para favorecer el aprendizaje de los estudiantes. Es decir, en un medio para incidir tanto en lo didáctico como en lo curricular. Y dejará de ser sólo una manera de constatar la incorporación o retención de determinados contenidos curriculares.

De acuerdo con Leymonié (2008), no es suficiente elaborar instrumentos de evaluación que simplemente constatar si los estudiantes han memorizado cierta información, es necesario elaborar actividades de enseñanza y evaluación que apunten a evaluar el grado de comprensión alcanzado, el razonamiento desarrollado y el uso de la información aprendida. En donde las actividades de evaluación sean a su vez actividades generadoras de enseñanza y las actividades de enseñanza ayuden a formular actividades de evaluación.

El gran desafío consiste entonces en hacer de la evaluación y de los exámenes a gran escala un problema didáctico que promueva no sólo que se aprenda, sino que se aprenda más y mejor.

## **2.2 Propuesta de un modelo de análisis de reactivos para los exámenes a gran escala**

La elaboración de un modelo didáctico para el análisis de reactivos de exámenes a gran escala, responde a la intención de devolverle el sentido didáctico a la evaluación a través de pruebas a gran escala. El propósito central consiste en procurar que a través de este modelo, la evaluación por medio de pruebas a gran escala proporcione información sobre el aprendizaje de los estudiantes y ayude a mejorar la enseñanza. Esto con el fin de que las pruebas estandarizadas y masivas no sólo verifiquen el aprendizaje, sino que lo promuevan e impacten en la enseñanza, porque evaluar es más que medir.

### **2.2.1 Consideraciones básicas sobre la evaluación del aprovechamiento escolar**

Hablar de evaluación del aprendizaje, es hablar de una práctica inherente a este proceso. Si se percibe la evaluación como un proceso que lleva a la comprensión de un fenómeno, en este caso, el aprendizaje (Díaz Barriga, 2005), es importante que la evaluación proporcione información sobre la realización, condiciones y obstáculos que se presentan en el acto de aprender<sup>55</sup>. La discusión sobre la evaluación del aprendizaje y sus implicaciones conceptuales y metodológicas, sobre todo, en relación a la elaboración de exámenes, permite ver de acuerdo con lo que Díaz Barriga (1997) ha venido planteado, la intención de que la evaluación proporcione un conocimiento sobre el proceso de aprendizaje de los seres humanos.

Por su parte Coll (2003) insiste en la necesidad de pensar en una evaluación orientada a constituir un eslabón entre las prácticas de enseñanza del profesor y las construcciones y adquisiciones de los alumnos en relación al aprendizaje<sup>56</sup>. Es decir, en un sentido formativo, la evaluación debe permitir influir en la enseñanza y ayudar a mejorar el aprendizaje. Así, aprendizaje, enseñanza y evaluación, tendrían que constituir una unidad indisoluble. En

---

<sup>55</sup> Díaz Barriga desde 1987, ha venido insistiendo en la ausencia de una construcción teórica de la evaluación en relación a su objeto de estudio, el aprendizaje, impide obtener una mejor comprensión de los procesos educativos en relación al aprendizaje (Díaz Barriga, 1987, 1997 y 2005a).

<sup>56</sup> Coll (2003) al abordar la función reguladora de la evaluación y el papel que juega en la construcción del conocimiento escolar, insiste en que las prácticas de evaluación deben ayudar a la toma de decisiones de orden pedagógico y didáctico por parte del profesor.

donde la evaluación esté al servicio del aprendizaje y la enseñanza<sup>57</sup> Mauri y Miras (1996) por su parte plantean la importancia de considerar los procesos de enseñanza y aprendizaje como el objeto principal de toda evaluación<sup>58</sup> Lo que implica que los instrumentos de evaluación den cuenta de los elementos que componen el aprendizaje. Es decir del tipo de contenido y nivel de conocimiento evaluado, de los procesos cognitivos que en éste intervienen, así como de los requerimientos en materia de enseñanza.

Según lo trabajado por algunos especialistas,<sup>59</sup> los procesos cognitivos se constituyen como el eje de la acción de enseñanza, de aprendizaje y por ende de evaluación, donde el contenido desempeña un papel importante, pues es el medio a través del cual se hacen visibles los procesos, pero no es lo fundamental ni de la enseñanza, ni del aprendizaje, ni de la evaluación. El acto de enseñar, aprender y evaluar, va más allá de que los alumnos retengan o ejerciten un determinado contenido, todo contenido, deberá ser trabajado y evaluado a través de ciertas operaciones cognitivas.

Desde esta perspectiva, todo instrumento de evaluación tendría entonces que recuperar el papel de las operaciones cognitivas en la enseñanza, en el aprendizaje y en la evaluación. Ante lo cual, los exámenes, deberían proporcionar información sobre el aprendizaje escolar, permitir la comprensión de elementos que pudieran mejorar la enseñanza y ayudar a la construcción de situaciones didácticas que favorezcan el aprendizaje, reconociendo las diferencias individuales y socioculturales que siempre están presentes en todo proceso educativo.

### **2.2.2 Elementos para una fundamentación conceptual del modelo que se propone.**

Con base en todo lo anterior y en vista de la ausencia de un análisis de reactivos para los exámenes a gran escala con enfoque didáctico, proponemos un modelo de análisis de reactivos que tiene el propósito de mostrar el sentido didáctico de los exámenes a gran escala.

---

<sup>57</sup> Sin embargo la proliferación y peso que los exámenes a gran escala han tenido en nuestro país, está ocasionando que más que evaluar lo que el alumno ha aprendido, alumnos y maestros deben "aprender" o por lo menos retener lo que viene en el examen. Con lo cual la relación didáctica se invierte.

<sup>58</sup> Para Mauri y Miras (1996), hablar de evaluación escolar, como se le llama a la evaluación realizada en los centros escolares en la escuela española, es hablar de los procesos de enseñanza-aprendizaje.

<sup>59</sup> En Uruguay a raíz del Programa PISA se ha realizado un trabajo muy interesante con docentes, donde se plantea la relación contenido-proceso en todo sistema de evaluación (véase Sánchez, 2008a y b).

La didáctica, como disciplina encargada de estudiar los procesos de enseñanza-aprendizaje, y el contexto donde se lleva a cabo, plantea la necesidad de identificar y organizar los elementos que aquí intervienen. Todo abordaje didáctico requiere considerar elementos de orden contextual, revelador de significados y ambientes socioculturales, aspectos lógicos del contenido, procesos cognitivos para su incorporación y aspectos metodológicos, que da cuenta de la aproximación al contenido

Así, todo aprendizaje requiere una organización lógica del contenido, una organización cognitiva para la activación de procesos, una organización metodológica que se refiere a la manera de llevarlo a cabo y un contexto sociocultural de ubicación y referencia

Porque aunque existen diferentes posturas en cuanto a la relación que existe entre contenido, proceso y contexto en una situación de aprendizaje, coincidimos con la que considera que la fuente del impacto educativo, está no en el contenido, sino en la actividad mental que el estudiante o el maestro trae a colaboración. (Taba, 1979, Aebli, 2001), lo que quiere decir que lo importante no es tanto el contenido, sino el impacto que éste ejerce sobre los procesos cognitivos

En palabras de Taba (1979)

“Cualquier materia puede ser igualmente eficaz o ineficaz, según los métodos de instrucción y aprendizaje que se empleen. Por ejemplo si el principal método es la memorización, no importa si se trata de informes históricos, fórmulas matemáticas o clasificación de plantas. Si el método es la solución de problemas y el análisis, casi cualquier tema puede producir impacto sobre el razonamiento crítico. un concepto como el de la justicia, tiene valor educativo sólo en la medida en que el alumno pueda hacer con él algo más que repetir su definición verbal” (p.231)

Con base en lo anterior fue que construimos la estructura didáctica del modelo de análisis que incluye elementos de orden lógico, referidos a los tipos y niveles de contenido, de orden cognitivo, referido a los procesos, de orden metodológico derivada de la tarea y referidas a las



---

experiencias de aprendizaje<sup>60</sup> y tipo de aprendizaje que se deriva y de orden sociocultural que incluye los contextos a los que se refiere.

La estructura didáctica del modelo que proponemos, está sustentado en una organización lógica, metodológica, cognitiva y sociocultural, que a continuación se describirá. El análisis de todos estos elementos en los ítems de las pruebas a gran escala, ayudará a hacer una traducción didáctica de los resultados, lo que le da un nuevo sentido a las tareas de evaluación de las pruebas a gran escala.

### **2.2.3 Descripción conceptual de un modelo didáctico y sus componentes, para el análisis de reactivos de pruebas a gran escala.**

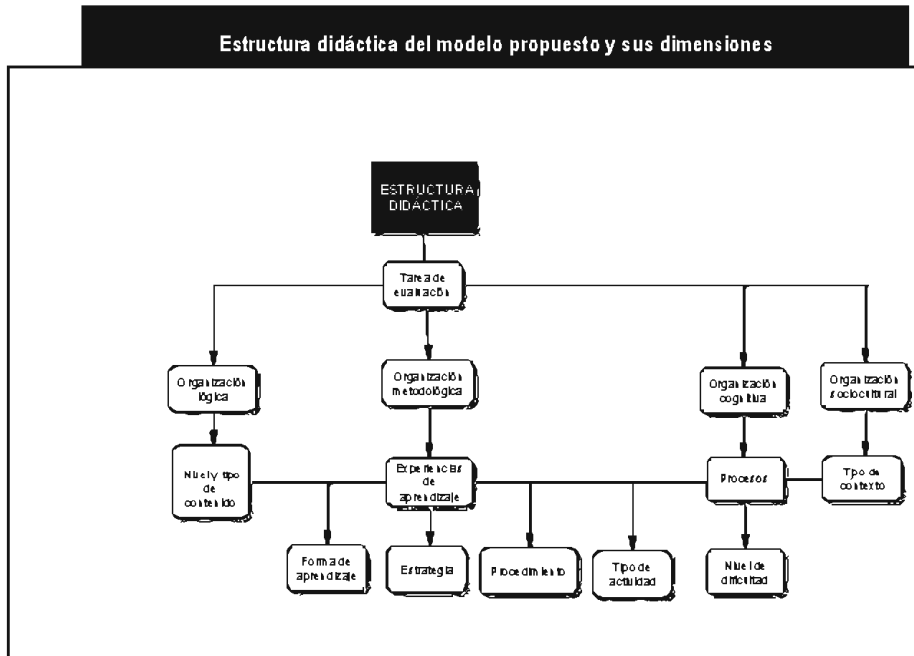
El modelo que aquí proponemos considera que para conocer las situaciones de aprendizaje y enseñanza implícitas en los ítems de una prueba, es importante analizar la estructura didáctica en su dimensión lógica, cognitiva, metodológica y sociocultural, en donde a partir de la tarea de evaluación se puede distinguir el nivel y tipo contenido al que se refiere, los procesos cognitivos implícitos, las tareas del aula, experiencias y tipo de aprendizaje que se derivan y los elementos de orden sociocultural que incluye. Porque de acuerdo con Taba (1979), todo proceso de aprendizaje no es más que la interacción entre el estudiante y un contenido de aprendizaje, de ahí la importancia de considerar en un análisis didáctico, tanto la organización y tipo de contenido, como la manera en que éste se hace llegar al estudiante. "Cada rama del conocimiento, tiene por lo menos dos características principales: su propia reserva de información adquirida y un método especializado para adquirir ese conocimiento" (p. 230).

La figura 1 muestra cómo quedó conformada la estructura didáctica del modelo que proponemos.

---

<sup>60</sup> La discusión sobre la organización de las experiencias de aprendizaje, es iniciada por Taba (1979).

Figura 1



Fuente: elaborado por la autora

Como se puede observar la estructura del modelo está compuesta por cuatro ejes, con diferentes dimensiones, en donde a partir de la tarea de evaluación, se analizará el nivel y tipo de contenido, las experiencias de aprendizaje que se derivan, los procesos implícitos y el tipo de contexto en el que se ubica. Cada una de estas dimensiones comprende diferentes elementos que tienen que ver con la enseñanza y el aprendizaje, en cuanto a su organización y contextualización. La organización lógica tiene que ver con la naturaleza del conocimiento. En la organización metodológica, las experiencias de aprendizaje establecen la relación con el conocimiento y la manera en como éste puede ser incorporado por los estudiantes. La organización cognitiva ayuda a identificar los procesos implícitos y el nivel de dificultad en el que se desarrolla. Y la organización sociocultural revela el significado que adquieren los contenidos en los diferentes

contextos o situaciones. Sin embargo, cabe mencionar que en la acción didáctica cada uno de estos ejes no se da por separado, sino que guardan una relación estrecha, de ahí la línea horizontal que une las diferentes dimensiones de los ejes.

El Cuadro 11 muestra la clasificación que se ha establecido en cada una de estas dimensiones y ejes.

Cuadro 11

**Clasificación de las dimensiones y ejes del modelo**

<b>Tarea de evaluación</b>		
Reproductivas de ejercitación rutinaria y poco significativas		
Problematizadoras y significativas que implican una comprensión y transferencia		
<b>Organización lógica del contenido</b>	Nivel y tipo de contenido	Hechos Ideas básicas o conceptos Procedimientos
<b>Organización metodológica</b>	Forma de aprendizaje	Significativo No significativo
	Estrategias	De recirculación de la información Elaboración, clasificación organización
	Tipo de actividades	Reproductivas Simbólicas Representativas Relacionadas con cosas Prácticas Reales Relacionadas con personas
<b>Organización cognitiva</b>	Procesos simples o complejos	<b>Simples:</b> observación reconocimiento identificación descripción recuperación de información ordenamiento clasificación <b>Complejos:</b> análisis síntesis Interpretación de textos Reflexión y evaluación
<b>Organización sociocultural</b>	Tipo de contexto	Disciplinar-escolares Extradisciplinar, vinculados a situaciones reales

A continuación describimos cada uno de los elementos que integran la estructura didáctica del modelo de análisis

- 1) Tarea de evaluación
- 2) Organización lógica del contenido en cuanto a su nivel y tipo
- 3) Organización metodológica
- 4) Organización cognitiva
- 5) Organización sociocultural

### 1) Tareas de evaluación

Las *tareas de evaluación*, para autores como Mauri (2007) consisten en las diferentes preguntas, ítems o problemas que los alumnos tendrán que responder, abordar o resolver en una situación de evaluación. Para su clasificación, tomamos los planteamientos que diversos autores hacen de su relación con procesos y situaciones de enseñanza-aprendizaje (Aebli, 2002, Snyders, 1972, Denyer, Furnémont, Poulain y Vanloubbeeck, 2007, Coll, 2003 y Meirieu 2004) <sup>61</sup> Con base en esto, rescatamos dos tipos de tareas: tareas reproductivas de ejercitación rutinaria y poco significativas que solicitan la reproducción de saberes aislados, (Aebli, 2001) centradas en ejercicios que requieren la práctica de algún contenido escolar (Snyders 1972, p.52) pero ajeno a la realidad y vida del niño. O por el contrario, tareas problematizadoras y significativas que implican una comprensión y transferencia en donde se plantea un enigma, a las que autores como Denyer, Furnémont, Poulain, Vanloubbeeck (2007), hacen alusión por su vinculación con el enfoque de competencias. Las cuales plantean la necesidad de establecer nexos y movilizar los conocimientos, de tal manera que el estudiante pueda encontrar la relación entre lo que aprendieron en la escuela y las situaciones con las que puede encontrarse en la vida (Meirieu, 2004).

El Cuadro 12 muestra los tipos de tareas que se consideran en el modelo junto con su definición y forma de identificarlos.

---

<sup>61</sup> Estos autores plantean diversos tipos de tareas asociadas a la enseñanza o al desarrollo de una competencia.

Cuadro 12

**Tipos de tareas para el análisis de los ítems**

Tipo de tarea	Definición	Indicadores
<b>Tareas reproductivas de ejercitación rutinaria y poco significativas</b>	Tareas que solicitan la reproducción de saberes aislados. Centradas en ejercicios que requieren la práctica de algún contenido escolar. Comprenden situaciones ajenas a la realidad y vida del niño.	Se pide definir conceptos y realizar ejercicios que implican sólo la práctica de algún contenido escolar en situaciones lejanas y sin sentido para el niño.
<b>Tareas problematizadoras que implican una comprensión y transferencia y resultan significativas</b>	Tareas que plantean un enigma exigen establecer nexos y aplicar su saber y saber hacer para la resolución de un problema. Permiten movilizar los conocimientos del alumno aplicándolos a una situación real. Comprenden situaciones cercanas a la realidad y vida del niño.	Son situaciones que requieren de nexos causales y finales para la solución de problemas que permiten a los alumnos movilizar sus conocimientos y establecer relaciones entre lo que aprendieron en la escuela y las situaciones con las que pueden encontrarse. Se solicitan causas y efectos de los fenómenos, motivos e intenciones. Ej. ¿Con qué consecuencias? ¿Para qué? Empleo de situaciones cercanas a la vida del niño.

Fuente: adaptado por la autora

**2) Organización lógica del contenido en cuanto a su nivel y tipo**

Pozo (2002), denomina resultados del aprendizaje a los contenidos. Éstos son fundamentales en toda tarea de evaluación y constituyen un importante componente del aprendizaje, por referirse a lo que se aprende. Constituyen la materia prima que moviliza la estructura cognitiva de los estudiantes.

Taba (1979), establece una *organización lógica del contenido* en cuanto a su nivel y tipo, que tiene que ver con la naturaleza del conocimiento, la diferenciación de los niveles de contenido y las distintas funciones que éstos cumplen, para lo cual establece diferentes niveles con respecto al conocimiento.

El primero nivel se refiere a contenidos relacionados con hechos específicos que requieren de un bajo nivel de abstracción. Pertenecen a esta categoría contenidos relacionados con descripciones, datos y hechos y enumeración de características. Obtener este tipo de contenido puede ser importante, aunque su dominio no genera ideas nuevas y tienen una alta probabilidad de pronto convertirse en obsoletos. Dewey (1993) se refirió a este tipo de conocimiento como "equipo e muerto". Dado que los hechos como tales, son sólo la materia prima, su función en el proceso de aprendizaje es efímera, no deben constituir la base fundamental del aprendizaje o de la evaluación.

El segundo nivel comprende contenidos conceptuales que requieren dotar de significado a la información o material presentado para comprenderlo (Coll, Pozo, Sarabia, y Valls, 1992). Incluyen interrelaciones entre elementos básicos y representan un nivel de abstracción mayor al conocimiento factual. Para Dewey (1993), son significados establecidos que comienzan con experiencias y se precisan y generalizan con el uso. Los conceptos o ideas básicas, representan otro nivel de conocimiento y se refieren relaciones causales, leyes científicas y principios matemáticos. En este nivel de contenido, podemos encontrar lo que se conoce como conceptos, definidos como "sistemas complejos de ideas altamente abstractas que sólo pueden estructurarse mediante experiencias sucesivas en una variedad de contextos" (Taba, 1979, p.236).

El tercer nivel es el referente a los procedimientos se refieren a saber hacer o saber actuar de manera eficaz. Tienen tres características básicas: acción, orden y meta, se trata de una forma de actuar sistemáticamente para la consecución de una meta (Coll, Pozo, Sarabia y Valls, 1992).

En algunos currículos se introducen las actitudes y valores como contenido educativo concreto. Sin embargo los valores y actitudes, constituyen parte de la experiencia misma y más que transmitirse, enseñarse o inclusive evaluarse, éstos, son parte de la vivencia diaria. De ahí que no constituyan en sí un contenido por separado, sino que son parte integrante de todas las materias de aprendizaje, por lo cual o pueden ser enseñados o evaluados de manera independiente, más bien se adquieren a través de la experiencia, como diría Dewey (1995-2004) para quien "el

punto a discurrir en una teoría de los valores educativos, es pues la unidad o integridad de la experiencia” (p 212) <sup>62</sup>

El Cuadro 13 presenta una descripción de los elementos considerados en la organización lógica del contenido

Cuadro 13

<b>Descripción de la organización lógica contenido</b>			
<b>Dimensiones</b>	<b>Clasificación</b>	<b>Definición</b>	<b>Indicadores</b>
Nivel y tipo de contenido	Hechos	Requieren un bajo nivel de abstracción. Se refieren a la adquisición de datos y hechos específicos.	Enumeración de características, identificación de símbolos, nombres y lugares.
	Conceptos o ideas básicas	Se refieren a relaciones que se pueden dar entre la información.	Implica clasificar para establecer categorías, principios, leyes y generalizaciones.
	Procedimientos	Se refieren a la acción sistematizada de saber cómo alcanzar un objetivo o meta.	Preguntas referidas al cómo se llegó a es decir al procedimiento que se siguió.

<sup>62</sup> En "Democracia y Educación" (1995-2004) Dewey desarrolla una teoría de la moral al plantear que el desarrollo moral se basa en que la escuela sea una comunidad de vida que permita el desarrollo de las percepciones y los intereses sociales de los alumnos en base a una experiencia común y con el objetivo de que esa experiencia debe continuarse fuera de la escuela es decir, debe haber continuidad entre el desarrollo moral social que se produce dentro y el de fuera de la escuela. No basta que el hombre sea bueno, tiene que ser bueno para algo, y la educación moral, forma parte de ese proceso de ser bueno para algo y desarrollar nuestra aportación a la sociedad en la que nos enmarcamos inevitablemente.

### 3) Organización metodológica

Las estrategias de aprendizaje pueden clasificarse en función del dominio de conocimiento y forma de aprendizaje que se quiere lograr y evaluar (Díaz Barriga Arceo y Hernández 2002, p. 238)

La organización metodológica del modelo que presentamos, tiene que ver con las experiencias de aprendizaje, ante lo cual, conviene considerar la relación que éstas guardan con la naturaleza del contenido y formas de aprendizaje, estrategias, procedimientos y tipos de actividades que se generan en el aula

Por ejemplo para el aprendizaje de un hecho específico, las estrategias de repetición y de repaso, se vuelven especialmente útiles cuando lo que se desea es acumular datos, para un aprendizaje repetitivo. Para un contenido conceptual a través del cual se pretende organizar, integrar y caracterizar información, la representación gráfica de redes conceptuales se considera lo más adecuado para un aprendizaje significativo (Díaz Barriga Arceo y Hernández, 2002)

Así, una tarea de evaluación, refiere un determinado contenido, que a ude a una forma de aprendizaje, requiere de estrategias y procedimientos específicos y comprende actividades en diferentes dimensiones

Para identificar las formas de aprendizajes, estrategias, y procedimientos implícitos en una tarea de evaluación tomamos como base la clasificación de Díaz Barriga Arceo y Hernández (2002) quienes presentan una clasificación de estrategias, vinculadas a una forma de aprendizaje que se promueve a través de distintos procedimientos. Así como también la taxonomía de actividades propuesta por Aebli (2001)

Entonces tenemos que ante un aprendizaje de tipo repetitivo, la estrategia básica es el repaso y la recirculación de la información, que se pueden realizar mediante procedimientos de repetición simple y copia. El repaso se vuelve una estrategia útil cuando los materiales que se han de aprender y evaluar presentan escasa significación lógica o psicológica para el estudiante. Estas estrategias son las más óptimas para



contenidos y aprendizajes memorísticos y repetitivos De lo que se trata es de repetir una y otra vez la información o contenido a aprender

Por el contrario un aprendizaje de tipo significativo, requiere de estrategias de elaboración y organización para ayudar a integrar y relacionar el nuevo contenido, con los contenidos previamente aprendidos Lo que requiere procedimientos de parafraseo, representación mental, identificación de palabras claves, elaboración de inferencias, analogías, redes semánticas, mapas conceptuales y uso de estructuras textuales

Para Aebli, (2001) "el aprendizaje se realiza en el proceso de su ejecución" (p. 20) y representa una determinada forma de actividad, ante lo cual presenta una taxonomía de la actividad y el saber con tres dimensiones que incluyen dos modos cada una

- 1) Las actividades relacionadas con cosas o con personas
- 2) Las actividades prácticas de tipo productivo o representativas, sólo de reconocimiento
- 3) Las actividades reales o las actividades simbólicas

Las actividades del aula que se derivan pueden estar relacionadas con cosas o con personas, o pueden estar encaminadas a producir algo, o simplemente a reconocer o representar una cosa o procedimiento También las actividades pueden manejar de manera directa una realidad o estar mediatizadas por signos o símbolos a través de lenguajes y medios figurativos Cada una de estas dimensiones favorece a su vez una determinada forma de aprendizaje

Cuadro 14

<b>Descripción de la organización metodológica</b>					
	<b>Dominio del conocimiento</b>	<b>Forma de aprendizaje</b>	<b>Estrategia</b>	<b>Procedimiento</b>	<b>Tipo de actividades</b>
Experiencias de aprendizaje	Factual	Repetitivo	Repaso Recirculación de la información	Repeticón simple Acumulación	Reproductivas
	Conceptual	Significativo	Selección Clasificación organización	Identificación de palabras claves Categorización Redes semánticas Parafraseo Inferencias	Simbólicas Prácticas reales Relacionadas con personas
	Procedimental	Repetitivo	Demostración	Reproducción	Representación

	Dominio del conocimiento	Forma de aprendizaje	Estrategia	Procedimiento	Tipo de actividades
		Significativo	Construcción	Elaboración	Producción Relacionadas con cosas o personas

Fuente adaptado de Díaz Barriga Arceo y Hernández (2002) y Aebí (2001)

#### 4) Organización cognitiva

Si consideramos que el contenido es la materia prima que moviliza la estructura cognitiva del alumno, es importante también atender a los procesos cognitivos implícitos en estos contenidos. Siguiendo a autores tanto del campo de la didáctica<sup>63</sup> como de la psicología cognitiva,<sup>64</sup> quienes han colocado a las operaciones cognitivas en el centro de su modelo de aprendizaje y evaluación, el modelo que proponemos, pretende la identificación de las *operaciones cognitivas* comprendidas en los reactivos, alrededor de las cuales se puede construir una acción didáctica. En donde las tareas de evaluación, permitan pensar en tareas del aula que desarrollen estos procesos.

Para el análisis de los procesos cognitivos construimos una clasificación obtenida de diferentes fuentes. En primer lugar recurrimos a la taxonomía utilizada por Castañeda (2006) para la construcción de exámenes a gran escala. Pero dado que sólo analizaremos los reactivos referentes a un área de conocimiento, que en este caso será la de español, es importante considerar que cada área de conocimiento, está vinculada a procesos específicos. Si consideramos que los reactivos analizados serán del área de Español, es necesario tomar en cuenta los diferentes tipos de procesos que en ésta intervienen. Los cuales van desde la decodificación, pasando por la identificación, hasta llegar a la integración semántica y a la interpretación (Mouchon, 2005). De ahí que consideramos conveniente tomar los niveles propuestos por la prueba PISA para la evaluación comprensión lectora,<sup>65</sup> además de la clasificación elaborada por Meza de Vernet (2002)<sup>66</sup> para el lenguaje.

<sup>63</sup> Entre estos podemos encontrar a Aebí, (2002) y Meirieu (1992)

<sup>64</sup> Nos referimos a Pozo (2002), quien concibe a los procesos cognitivos, como los mecanismos mediante los cuales se produce el aprendizaje y Castañeda (2006)

<sup>65</sup> PISA presenta una definición de comprensión lectora (OCDE, 2001), a partir de la cual, establece como una de las dimensiones para su evaluación, los procesos, entendidos como las tareas de lectura que los estudiantes necesitan realizar cuando se enfrentan a un texto (NEE, 2008, p. 31). Tenemos entonces que la clasificación

El Cuadro 15 presenta y describe los procesos utilizados en esta clasificación

**Cuadro 15**

**Procesos cognitivos para el área de español**

Procesos simples	Observación	Consiste en focalizar la atención en un objeto o situación para identificar sus características, de acuerdo con su interés o con un objetivo previamente definido
	Reconocimiento	Es uno de los procesos más elementales y está asociado con la memoria sensorial, constituye por procesos fisiológicos simples (Bourne, Dominowski, Loftus y Healy, 1986)
	Identificación	Demanda reconocer información (Castañeda, 2006, p 155)
	Ordenamiento	Consiste en establecer una secuencia entre objetos, hechos, seres, etc., basándose en la identificación de las características que cambian (Meza de Vernet, 2002)
	Descripción	Permite enumerar e integrar las características de lo observado, en un todo significativo. Es el resultado de la observación sistemática (Meza de Vernet, 2002)
	Recuperación de la información	Implica localizar información relevante de los aislados o vinculados con otros elementos del texto (INEE, 2008)
	Comparación	Consiste en establecer semejanzas y diferencias entre dos o más objetos, hechos o situaciones a partir de variables (Meza de Vernet, 2002)
	Clasificación	Consiste en reunir objetos o conceptos en grupos llamados clases, de acuerdo con un criterio. Los grupos tienen dos condiciones: son mutuamente excluyentes y no omiten ningún representante del grupo (Meza de Vernet, 2002)
Procesos complejos	Análisis	Implica dividir un todo en las partes y relacionar parte-todo y todo-parte (Castañeda, 2006, p. 157)
	Síntesis	Es complementario al análisis y consiste en integrar el todo en una nueva unidad con significado (Meza de Vernet, 2002)
	Interpretación de textos	Se refiere a encontrar el significado y llegar a alguna conclusión a partir del texto escrito (INEE, 2008)
	Reflexión y evaluación	Implica relacionar la información escrita en cuanto a forma y contenido, con sus conocimientos, ideas y experiencias anteriores (INEE, 2008)

establecida por PISA para la evaluación de la comprensión lectora, aporta elementos importantes para el análisis de reactivos del área de español

<sup>66</sup> Especialista en procesos de aprendizaje y coordinadora de procesos de lenguajes de la Universidad Metropolitana de Venezuela

El tipo de procesos implícitos en los reactivos, permitirá a su vez ubicar el nivel de dificultad

### 5) Organización sociocultural

La situación o contexto en donde son presentadas las tareas de un ítem, forma parte de la situación sociocultural. Para su definición, tomamos como base los contextos del marco conceptual de PISA (OCDE, 2006) y la síntesis que de éstos se ha realizado en algunos países para definir las actividades de una prueba<sup>67</sup> (ANEP, 2007a). Con base en esto se definen dos tipos de situaciones o contextos donde pueden ser ubicadas las tareas de evaluación: El disciplinar de tipo escolar, relacionado con contenidos disciplinares y cuestiones escolares y el extradisciplinar, vinculado a situaciones reales.

**Cuadro 16**

**Tipo de contextos**

Tipo de contexto	Definición
Disciplinar de tipo escolar	Relacionado con contenidos disciplinares y cuestiones escolares.
Extradisciplinar, vinculados a situaciones reales	Cuestiones situadas dentro de un contexto real, relacionadas con aspectos de su entorno. Éstas pueden ser de tipo personal u ocupacional.

Fuente: adaptado de ANEP (2007a) *Programa PISA Uruguay* CODICEN-DIE

En el desarrollo del presente capítulo hemos visto como tanto la construcción del reactivo como su contenido son elementos importantes a considerar para el análisis de los mismos y las inferencias que de éste se derivan. Surge entonces la necesidad de conocer no sólo su estructura didáctica del reactivo sino también de identificar algunas características con respecto a su construcción que afecten su resolución.

<sup>67</sup> Nos referimos a la experiencia de Uruguay, en especial de la Administración Nacional de Educación Pública (ANEP), quienes a partir de la participación de Uruguay en PISA, han elaborado unos boletines de difusión denominados *PISA en Uruguay*, para dar a conocer a docentes y público en general los marcos conceptuales y elementos más importantes de la prueba PISA, con el fin de aprovechar lo que el proyecto PISA ofrece para a partir de una reflexión colectiva con docentes, ayudarles a mejorar sus prácticas sobre la base de lo PISA significa (Sánchez, 2008).

## 2.2.4 Variables de construcción del reactivo que afectan su resolución

El análisis de los reactivos de una prueba, no puede desconocer ciertos elementos, relacionados con la construcción del reactivo que afectan su resolución

Con la intención de complementar el modelo de análisis didáctico de reactivos para pruebas a gran escala, a partir de los lineamientos enunciados al inicio de este capítulo, hemos identificado cinco elementos que también son importantes a considerar con respecto a la redacción de un reactivo

### 1) **Objetivo de evaluación de ítem**

Especialistas en la materia (Sánchez, 2008) lo definen como lo que el ítem pretende evaluar. En donde el enunciado puede combinar contenido y proceso cognitivo

Los reactivos, deben responder principalmente al objetivo para el cual fue creado, así como al dominio de referencia, a la población de sujetos a evaluar y a las condiciones para su aplicación

### 2) **Expresiones socioculturales del lenguaje**

De acuerdo con Tort (1979), las condiciones de un test no son neutras. "Es una situación de clase que favorece desigualmente el éxito de los niños según su origen social. Las preguntas que se formulan, no tienen el mismo sentido para las distintas clases sociales (Tort, 1979, p.49)

Bernstein (1985) al estudiar las condiciones sociales que influyen en el aprendizaje escolar y el significado que adquieren los códigos en diferentes contextos, ha empleado el concepto de *código sociolingüístico* para referir "la estructuración social de los significados" y sus manifestaciones lingüísticas en contextos diferentes pero relacionados" (Gómez y Hernández, 1991, p. 286). Esto con el propósito de explorar

como "los sistemas simbólicos" son manifestaciones reguladoras de la estructura de las relaciones sociales

### **3) Características perceptuales y de redacción de los distractores**

La teoría de la "Gestalt", considerada también como "teoría de la estructura" o "teoría de la configuración o forma", es la encargada de estudiar la incidencia en los sistemas totales, en las estructuras en las que las partes están interrelacionadas dinámicamente de manera que el todo no puede ser inferido de las partes consideradas separadamente. Una de las contribuciones más importantes de la Gestalt, reside en las aportaciones al estudio de la percepción, al definir una serie de principios de organización perceptiva que permiten captar de forma integral estas totalidades o *gestalts*. Uno de estos principios, la *Ley de la igualdad o equivalencia*, plantea que cuando concurren varios elementos con los que son iguales, hay una tendencia a constituir grupos con los que son iguales. Si se potencian las formas iguales se establecen condicionantes potenciadores para el fenómeno agrupador de la percepción visual. Al considerar que la percepción visual es un proceso donde se crean patrones de significados a partir de la información sensorial pura, ésta se convierte en uno de los condicionantes que influyen directamente en la visualización de la información.

El papel de un distractor dentro de la teoría del test, es impedir dar con la respuesta correcta cuando no se sabe. Por lo que es importante considerar las normas establecidas por los expertos para su redacción, de las cuales ya hemos hecho mención. Es fundamental cuidar que éstos no tengan dificultades perceptuales, estén bien escritos, sean lógicos y tengan sentido, tanto en sí mismos como con respecto a la base, para impedir que se llegue a la respuesta correcta por eliminación y realmente se cumpla el objetivo del ítem de detectar el aprendizaje y no de hacer suficientemente confusos los distractores.

### **4) Estrategias que los estudiantes emplean para su resolución**

No sabemos de estudios empíricos que muestren las diversas estrategias que pueden utilizar los estudiantes para la resolución de la pregunta. Sin embargo es factible que muchas de éstas se realicen por proceso de eliminación cuando los distractores no estén bien redactados, lo que indica la necesidad de investigar sobre tales procesos.

Hasta el momento, lo único que encontramos son algunas recomendaciones para contestar pruebas estandarizadas y éstas se presentan con la única intención de preparar al estudiante para el examen<sup>68</sup>. A raíz de las consecuencias que los resultados de las pruebas tienen para los docentes,<sup>69</sup> el preparar para el examen de manera directa, se ha vuelto una de las principales preocupaciones de los docentes, escuelas y sociedad en general. Mehrens (1989), en su artículo analiza las distintas formas en que se puede preparar a los estudiantes para la prueba y las implicaciones de estos "entrenamientos"

### 5) Justificación de distractores

Se trata de identificar si los distractores reportan un determinado error en el manejo del contenido. Esto, con el fin de que una vez detectados estos errores se describan sus características, dejando abierta la posibilidad de estudiar sus causas. Mismas que pueden ser intrínsecas a la evaluación (grado de dificultad del instrumento aplicado o adecuación a los contenidos abordados al desarrollo cognitivo de los estudiantes, entre otros) extrínsecas (características del currículo, estrategias de enseñanza aplicadas, etc.) (Sánchez, 2001)

Esto puede llevar a un análisis de errores persistentes y permitir así dar una más exacta retroalimentación tanto al estudiante como al docente

Así, los elementos del modelo que serán considerados para el análisis de los reactivos de español de la prueba ENLACE 2008 son los siguientes

---

<sup>68</sup> En textos publicados por Santillana, con el nombre de "Aprobado", se les da a los estudiantes diversas recomendaciones para "aprobar" los exámenes. Entre éstas se encuentran "Rellenar con tu lápiz el círculo correspondiente de la respuesta correcta cuidando de que esté completamente cubierto," porque un dispositivo electrónico calificará los exámenes y si los círculos no están completamente sombreados, el mecanismo considera que no contestaste" (Moreno, Huerta y Castellanos, 2007)

<sup>69</sup> Una de las más recientes declaraciones oficiales, como parte de la llamada "Alianza por la calidad de la educación" consiste en que los maestros cuyos alumnos mejoren los resultados de ENLACE, recibirán un bono adicional de bonificación. De acuerdo con los anexos técnicos de la Alianza por la Calidad de la Educación, habrá 6 tipos de bono, designados con base en el salario mínimo mensual. Para la asignación de éste, se tomará en cuenta los resultados de la prueba ENLACE del ciclo escolar anterior. El bono se entregará anualmente a los maestros y será de hasta 5 salarios mínimos dependiendo de los resultados que se obtengan. Del Valle (2008). "Prevén que el bono mejore a maestros" *Reforma* 17 mayo 2008 <http://www.reforma.com/nacional/articulo/442882647/>

**Cuadro17**

**Elementos del modelo considerados para el análisis**

Aspectos	Dimensiones
Estructura didáctica	Tipo de tarea de evaluación
	Nivel de contenido
	Forma de aprendizaje
	Estrategias y procedimientos
	Tipo de actividades
	Procesos
Variables de construcción del reactivo que afectan su resolución	Tipo de contexto
	Objetivo del ítem
	Expresiones socioculturales del lenguaje
	Características preceptuales y de redacción de los distractores
	Estrategias que los estudiantes emplean para su resolución
	Justificación de distractores

Fuente. elaborado por la autora

Los elementos presentados en este capítulo para un análisis didáctico de los reactivos de pruebas a gran escala, serán la base para el análisis de los reactivos de Español de la prueba ENLACE 2008, objeto de estudio de esta investigación



### CAPÍTULO 3

#### LA EVALUACIÓN DE LA COMPRENSIÓN LECTORA EN LA PRUEBA ENLACE DESDE UN ANÁLISIS DIDÁCTICO

*Los “números” sólo pueden tener algún significado si uno comprende qué tipo de aprendizajes fueron evaluados y desde qué concepción (Rave la).*

El propósito de este capítulo es aproximarnos al contenido y sentido evaluativo que adquiere una prueba masiva que se aplica a nivel nacional denominada como Evaluación Nacional de Logro Académico en Centros Escolares, mejor conocida como ENLACE<sup>70</sup>

Presentamos en un primer término una caracterización general de esta prueba, conforme a la información obtenida<sup>71</sup> sobre su proceso de elaboración y algunas reflexiones que surgen al respecto

Posteriormente con base en el modelo propuesto, analizamos algunos de los reactivos de español de la prueba de 6° grado del 2008. Finalmente, a partir del análisis didáctico de los reactivos seleccionados, ofrecemos una caracterización general de la evaluación de la comprensión lectora en la prueba ENLACE y una aproximación a los alcances, límites y sentidos encontrados a esta prueba. Esto a partir de lo que se dice y no se dice, pero que se puede decir, de esta prueba

---

<sup>70</sup> Esta prueba se aplica en todas las escuelas de educación básica del país. En su primera aplicación tuvo una cobertura de cerca del 90 por ciento, aplicándose a ocho millones de alumnos de tercero, cuarto, quinto y sexto grado de primaria y también de tercero de secundaria. En el 2009, se aplicó a más de 13 millones de estudiantes de educación básica en México y su cobertura se amplió, al aplicarse a alumnos de primero y segundo grado de secundaria (consultado en <http://www.enlace.sep.gob.mx/tr/>, diciembre 2009). Todo esto nos habla de la magnitud que ha alcanzado la prueba. El hecho de que se aplique en más de 100 mil escuelas del país de corte público y privado, urbanas y rurales, así como a las que atienden poblaciones en situaciones vulnerables como la población indígena, con los cursos comunitarios del CONAFE y las telesecundarias (<http://www.presidencia.gob.mx/prensa/dicursos/?contenido=28926>), revela que estamos ante una prueba masiva aplicada de manera uniforme a una diversidad de contextos escolares con características socioeconómicas y culturales distintas. Esto nos lleva a cuestionar la validez de aplicar una prueba homogénea, a poblaciones heterogéneas.

<sup>71</sup> Cabe destacar que mucha de la información que se tiene de ENLACE, se obtuvo de manera confidencial, sobre todo en cuanto a su creación y proceso de elaboración y validación. Pues esta información no es de acceso público.

### 3.1. Hacia una caracterización de la prueba ENLACE

Para caracterizar a la prueba ENLACE, es importante conocer sus antecedentes, elementos y concepciones que la definen, así como sus procesos de elaboración, selección, tipificación de reactivos y mecanismos de aplicación. Esto permitirá tener una idea de sus particularidades y propósitos a los que responde su elaboración y aplicación.

#### 3.1.1 Antecedentes y origen

ENLACE es una prueba a gran escala<sup>72</sup> elaborada y aplicada por personal de la Dirección General de Evaluación de Políticas (DEGEP) de la SEP. La gestación de esta prueba y su denominación es de naturaleza política.<sup>73</sup> Sus antecedentes los podemos encontrar en la década de los noventa con la prueba *Evaluación de la Educación Primaria (EVEP)*, prueba elaborada como parte del programa para abatir el rezago educativo.

Esta prueba cuenta con el asesoramiento técnico de diversas instituciones como

- Universidad Complutense de Madrid
- Sociedad Matemática Mexicana
- Centro de Investigaciones en Matemáticas, A. C.
- Instituto de Evaluación e Ingeniería Avanzada, S. C.

Esto da una idea que su surgimiento y aplicación obedece a intenciones más de tipo político que educativo y su realización está más centrada en aspectos técnicos que pedagógicos.

---

<sup>72</sup> Hablar de pruebas a gran escala es hablar de instrumentos experimentales estandarizados compuestos por un conjunto de tareas organizadas bajo condiciones uniformes. Los cuales son aplicados a un gran número de personas. Su finalidad es comparar y clasificar al sujeto de manera cuantitativa o tipológica (Benedito, 1986). Cabría preguntarnos entonces si realmente esta prueba es aplicada en condiciones uniformes ante la diversidad de contextos escolares con los que contamos en México.

<sup>73</sup> Nace cuando a raíz de un conflicto interno, en la Dirección General de Evaluación de la SEF, llega a la Dirección de Evaluación y Análisis del Proceso Educativo, de la SEP, un asesor externo de la Universidad Complutense de Madrid, el Dr. José Luis Gavina quien a partir de la existencia de una prueba anterior a la de ENLACE, denominada *Evaluación Diagnóstica Censal*, propone que se le denomine mejor *Evaluación Nacional de Logro Académico en Centros Escolares (ENLACE)*. Nótese como la denominación de Centros Escolares es de origen español, no es una denominación que utilizemos frecuentemente en México. En México se dice vamos a la escuela, no al Centro Escolar. Esto hace pensar en qué tanto la gestación de ENLACE estuvo pensada para el contexto mexicano.

### 3.1.2 Elementos que la caracterizan

La prueba de *Evaluación Nacional del Logro Académico en Centros Escolares*, mejor conocida como ENLACE, se define como una prueba del Sistema Educativo Nacional que se aplica anualmente<sup>74</sup> a los planteles públicos y privados de todo el país y que presenta las siguientes características

- Prueba censal<sup>75</sup> con la que se pretende evaluar el logro académico en las asignaturas de español y matemáticas<sup>76</sup> de los alumnos de 3<sup>o</sup> a 6<sup>o</sup> grado de educación primaria y de 3<sup>o</sup> grado de educación secundaria
- Prueba estandarizada, considerando que su administración se hace en condiciones “iguales” para todos los examinados
- Prueba objetiva porque su metodología de calificación es precisa y preestablecida, para proporcionar una referencia de comparación nacional
- Prueba de opción múltiple, donde la respuesta correcta es una sola. Cuenta con un procesamiento automatizado de la información a través de lector óptico
- Se conforma por un cuadernillo de preguntas y una hoja de respuestas, donde el alumno codifica sus respuestas a papel y lápiz

<sup>74</sup> Hasta la fecha lleva cinco aplicaciones en nuestro país: en el 2006 se aplicó del 5 al 9 de junio; 2006, en 2007, comenzó el 23 de abril; en el 2008 fue del 14 al 18 de abril y en el 2009 del 23 al 29 de abril, semana a la que se le ha denominado a partir del 2008, como la *Semana nacional de evaluación*. Los resultados de la aplicación 2009 en educación básica, fueron publicados en la página <http://www.enlace.sep.gob.mx/br/?p=n01> a partir del 5 de octubre del 2009. Esto nos habla de un espacio de seis meses para procesar la información y otros seis meses para preparar la siguiente aplicación. Ante lo cual cabe preguntarse si seis meses, serán suficientes para conocer y trabajar en los errores detectados. Al parecer estamos ante una acción evaluativa, que tiene la intención de solo medir por medir, sin dar el suficiente tiempo para analizar los errores y usarlos como medio para aprender. Esto no ha sido contemplado aún, como parte de los procesos de elaboración y aplicación de esta prueba.

<sup>75</sup> Se conoce como prueba censal aquella prueba que es aplicada a todos los estudiantes y en todas las instituciones educativas de los grados seleccionados. Mientras que una prueba muestral se selecciona a un conjunto de estudiantes de una población (objetivo). La cobertura de esta prueba ha ido aumentando. Según datos de la SEP, mientras en el 2006 se aplicó a 9,529,490 alumnos en 112,912 escuelas, para 2010, alcanzó una cobertura del 15,766,608 alumnos en 133,497 escuelas. <http://enlace.sep.gob.mx/ba/?p=aplica>. Consultado mayo 2010.

<sup>76</sup> Para el año 2008 se agregó una tercera materia que fue la de ciencias naturales. Esta tercer asignatura se irá rotando cada año para permitir una evaluación integral. La programación quedó de la siguiente manera: para 2009, se incorporó formación cívica y ética y para el 2010, se pretende incorporar historia, en 2011, geografía y en 2012, volver con ciencias. También en el 2008 se inició con la aplicación de ENLACE a la Educación Media Superior. Ante la incorporación en el 2009 de formación cívica y ética como área de evaluación, cabe preguntarse si la formación cívica y ética es algo que se pueda evaluar a través de una prueba, o es algo que se vive y que se vive en las actitudes diarias y que difícilmente puede ser medido.

- Prueba centrada en el conocimiento ya que evalúan el resultado del trabajo escolar contenido en los planes y programas de estudio oficiales vigentes. Está centrada en contenidos curriculares
- De aplicación autoadministrable y controlada
- Está compuesta por 50 reactivos como mínimo y 74 como máximo para cada asignatura
- Proporciona un diagnóstico individual por grado y asignatura
- Son de aplicación masiva.

### **3.1.3 Mecanismos de aplicación**

La aplicación de la prueba ENLACE, es coordinada por la Dirección General de Evaluación de Políticas de la Secretaría de Educación Pública. En cada entidad federativa es organizada por el área o instituto estatal de evaluación.

La impresión de la prueba, empaque y distribución nacional se realiza a través de la Comisión Nacional de Libros de Texto Gratuito.

Para el 2008, participan en la operación, además de los alumnos, un coordinador externo de aplicación, el director del plantel y los docentes-aplicadores. El director deberá invitar a padres de familia como supervisores, así como a personal externo a la institución como observadores.

El coordinador general de aplicación será el titular de las áreas estatales de evaluación o institutos estatales de evaluación quienes se auxiliarán de coordinadores regionales así como de la estructura educativa de primarias y secundarias.

El coordinador de aplicación trasladará los materiales de aplicación, normativos y de apoyo a la escuela designada. Se acreditará con el director al inicio de las actividades escolares y explicará en detalle la logística de aplicación. Verificará la presencia de padres-supervisores y observadores externos, así como del intercambio de grupos entre los docentes-aplicadores. Es el encargado de levantar las actas de inicio, cierre y en su caso, de irregularidades.

Se contrata además un coordinador externo de aplicación por cada plantel educativo, que en coordinación con el director de la escuela, serán responsables del proceso en

dicho centro escolar. La SEP comisionó en el 2008 aproximadamente a 133 mil coordinadores externos de aplicación, uno por cada plantel educativo, quienes son responsables de trasladar los materiales hasta el plantel y, en coordinación con el director de la escuela, garantizan la aplicación y levantamiento de las actas previas

El día de la aplicación, el coordinador externo acude a la sede con los exámenes empacados y resguardados en bolsas cerradas con sellos que garantizan su confidencialidad

El coordinador externo de aplicación se acredita con el director al inicio de las actividades, explica la logística y verifica la presencia de padres-supervisores, observadores externos y el intercambio de grupos entre los docentes-aplicadores

Durante la aplicación de la prueba, cada alumno recibe un cuadernillo impreso, constituido por reactivos de opción múltiple y una hoja de respuestas con un talón desprendible, el cual se entrega a cada alumno cuando finalizan la prueba como un comprobante de su participación que les servirá para consultar sus resultados. Las respuestas deberán ser anotadas, en una hoja de respuestas, diseñada para ser leída por un lector óptico

La aplicación de la prueba se realiza durante dos días y consta de cuatro sesiones de 45 minutos cada una, con un intervalo de 15 minutos de descanso. Durante este lapso los alumnos no conservan las pruebas ni hojas de respuesta

Las pruebas incorporan al final, preguntas para identificar a estudiantes de origen indígena y recabar información sobre su dominio del español y su comprensión de las preguntas realizadas en el examen

El proceso de aplicación es supervisado por aproximadamente 3 millones de padres de familia y observadores ciudadanos, y se realiza sin suspender labores ni autorizar comisiones a los maestros

Adicionalmente a las pruebas que se aplican a los alumnos, los directores del plantel contestan un cuestionario de contexto y, en los planteles seleccionados como "escuelas muestra", alumnos, maestros y padres de familia responden otro, que permite hacer un mejor análisis de los resultados y contribuye a su comparabilidad.

Los resultados son filtrados por un software de “detección de probabilidad de copia” que utiliza los métodos K-index y Scrutiny, que tienen como base patrones de respuestas incorrectas similares. En caso de detectar alguna irregularidad se registra en los reportes individuales para el alumno.

Los directores escolares cuidarán no suspender labores durante esos días ni autorizar ninguna comisión a los maestros de su plantel.

Además para el 2008, se contó con documentos normativos para la aplicación, entre los que se encuentran los siguientes:<sup>77</sup>

- Normas operativas
- Características generales del levantamiento de datos
- Manual para el coordinador operativo y/o regional
- Manual para el coordinador de aplicación
- Guía para el docente-aplicador
- Guía para el docente-aplicador de escuelas multigrado y del consejo nacional de Fomento Educativo (CONAFE)
- Guía para el docente-aplicador en escuelas de educación indígena
- Guía para padres supervisores

### 3.1.4 Cobertura

El Cuadro 18 muestra la cobertura de aplicación de la prueba ENLACE desde su primera ocasión, hasta el 2010. Como podemos observar, ésta ha sido variada tanto en escuelas como en alumnos. Sin embargo, en cuanto a alumnos se ha mantenido en un promedio de 10 millones. Ante lo cual cabe destacar la inversión que eso implica.<sup>78</sup>

<sup>77</sup> Véase <http://enlace.sep.gob.mx/ba/?p=docs>. Consultado en diciembre del 2009.

<sup>78</sup> En el 2008, esta inversión fue superior a 220 millones de pesos. Jorge Santibáñez, titular de la Unidad de Planeación y Evaluación de Políticas Educativas, informó que el costo de dicho examen es de 17 pesos por alumno. Consultado en: <http://www.vanguardia.com.mx/diario/noticia/politica/razonable-se-aplicara-la-prueba-enlace-en-escuelas-primarias-sep/149121>, diciembre, 2009.

Cuadro 18

**Cobertura de la prueba ENLACE 2006-2009**

Año	Escuelas	Alumnos	Asignaturas evaluadas
2010	133,497	15,766,608	-Español -Matemáticas -Historia
2009	129,669	13,187,688	-Español -Matemáticas -Formación cívica y ética
2008	121,668	9,930,309	-Español -Matemáticas -Ciencias
2007	121,585	10,148,666	-Español -Matemáticas
2006	112,912	9,529,490	-Español -Matemáticas

Fuente tomado de <http://enlace.sep.gob.mx/ba/?p=aplica>, consultado en mayo 2010

El cuadro 19, muestra de manera específica la diversidad de contextos de aplicación que tuvo en el 2008. Cabe señalar la heterogeneidad existente en estos contextos, que indudablemente se contraponen a la homogenización existente de la elaboración de pruebas estandarizadas para validar el aprendizaje.

Cuadro 19

**Contextos de aplicación de la prueba ENLACE 2008**

**PRIMARIA**

Escuelas ENLACE 2008	CONAFE	GENERAL	INDIGENA	PARTICULAR	TOTAL
Nacional	11,611	67,921	9,757	7,400	96,689

**SECUNDARIA**

Escuelas ENLACE 2008	GENERAL	PARTICULAR	TELESECUNDARIA	TECNICA	TOTAL
Nacional	6,965	3,772	12,210	3,957	31,904

Fuente: DGEP. 2008

### 3.1.5 Estrategia de operación

Cabe destacar que un acierto importante de las últimas aplicaciones de ENLACE es que ya se cuenta con mayor información en cuanto a su proceso de elaboración y aplicación. Las normas establecidas para tal efecto, hacen pensar que es una acción cuyas normas de operación están muy cuidadas. Sin embargo cabría preguntarse, si en un proceso evaluativo que pretenda conocer y mejorar la situación de aprendizaje de los estudiantes, ¿que es lo más importante cuidar y establecer mecanismos claros de normatividad? o cuidar el uso y análisis de la información obtenida. Si bien es cierto que la evaluación educativa debe ser un proceso sistemático y cuidado, es importante que para que este pueda convertirse en un importante motor de cambio, (Ravela, 2007) deberá cuidarse también el uso de los resultados de las evaluaciones, de tal manera que esté encaminado a orientar el proceso tanto de enseñanza como de aprendizaje. Es decir, deberá buscarse la manera en que tanto docentes como estudiantes, conozcan no sólo los resultados obtenidos sino el porqué de los mismos y la manera en que puedan mejorarlos.

### 3.1.6 Proceso de elaboración y construcción de la prueba

La construcción de pruebas a gran escala a nivel nacional, según documentos internos del Instituto Nacional de Evaluación (INEE), pasa por seis etapas básicas, cada una de las cuales comprende diversos procedimientos (Contreras, Backhoff y Larrazolo, 2004). Estas etapas son:

- 1) Definición del dominio de resultados que pretende el currículo. Se dice que por ser una prueba a gran escala alineada al currículo, una de las primeras etapas es hacer explícitas las intenciones de formación que plantea el currículo.
- 2) Análisis del currículo. Aquí lo que se pretende es identificar a través de distintas fuentes, el dominio curricular sobre el cual se desarrollará el examen y los resultados importantes a evaluar. Con base en el dominio definido previamente, se efectúa un análisis curricular del mismo a fin de estructurar los resultados de aprendizaje importantes que pretende el



---

currículum del área. En un test criterial, la validez del examen, dependerá de qué tanto se completen las metas de aprendizaje del currículum. Esto se realiza generalmente a través de la elaboración de una retícula de contenidos a evaluar, que no es más que una red que permite mostrar los contenidos y las relaciones que existe entre éstos y los cuales permite identificar. Contenidos fuente, prestan servicio a otros contenidos), contenidos sintéticos, (los que reciben servicios de otros contenidos), contenidos rama o de enlace, (los que dan y reciben servicios simultáneamente) y contenidos aislados (que no presentan relación con otros contenidos) (Contreras, Backhoff y Larrazolo, 2004)

- 3) Desarrollo del plan de evaluación. Se refiere al plan para construir la prueba, el cual incluye un muestreo representativo de los resultados de aprendizaje a evaluar ante la imposibilidad de medir todo el dominio de resultados que pretende el currículum, así como el diseño de las especificaciones del ítem, tanto desde el punto de vista del contenido, como del psicométrico y una capacitación a las personas que elaboraran los ítems.
- 4) Producción y validación de ítems. Consiste en un proceso de varios pasos: Elaboración del reactivo, análisis del reactivo a la luz de las especificaciones dadas, prueba del reactivo, análisis de las respuestas obtenidas mediante su aplicación, revisión de los ítems y estructuración de la versión final.
- 5) Análisis primario. Consiste en revisar la calidad técnica del ítem y en diseñar los procedimientos para calificar el examen y analizar los resultados.
- 6) Análisis secundario. Se trata de la elaboración de reportes de resultados.

El Cuadro 20, presenta una síntesis de estas etapas con sus procedimientos correspondientes.

Cuadro 20

**Etapas para la elaboración de pruebas**

1	Definir el dominio de resultados que se pretende	Selección y capacitación del Comité diseñador del examen
		Análisis del contenido curricular
		Análisis complementario (Colaboración de profesores en servicio)
2	Análisis del currículo	Elaboración de la retícula (redes) del contenido a evaluar
3	Desarrollo de un plan de evaluación	Muestreo de resultados de aprendizaje a evaluar
		Diseño de especificaciones de ítems
		Capacitación del Comité elaborador de ítems
4	Producción y validación de ítems	Elaboración de ítems según las especificaciones
		Revisión de la congruencia ítem-especificación
		Ensayo empírico y revisión de ítems
		Revisión de la prueba y estructuración de una muestra de ítems representativa del dominio curricular
5	Análisis primario	Ensayo empírico de gran escala y revisión de ítems
		Especificación de estándares de calidad para el contenido del ítem, para la calidad técnica del ítem y para la calidad integral de los puntajes del examen
6	Análisis secundario	Análisis de resultados para elaborar los reportes de resultados por escuela y municipio, por eje temático del plan de estudios y por destinatario
		Análisis de resultados para elaborar los reportes de resultados por escuela y municipio, por eje temático del plan de estudios y por destinatario

Fuente: tomado de Contreras, Backhoff y Larrazolo, 2004, p. 3

Un aspecto importante a considerar también en la elaboración de pruebas es el tiempo asignado a cada una de estas etapas. El Cuadro 21, presenta el calendario de elaboración que se siguió para el caso de la prueba ENLACE, 2006.<sup>79</sup>

<sup>79</sup> Presentamos el calendario de elaboración de la prueba ENLACE 2006, porque no contamos con el calendario del 2008, pero en sí, el tiempo transcurrido en la aplicación y evaluación es el mismo en cada aplicación.

Cuadro 21

**Calendario de elaboración para la prueba ENLACE 2006**

Actividad	Fecha	Observaciones
Desarrollo y definición de matrices de especificaciones.	Febrero	Tarea realizada por la SEB
Jueceo y Piloteo	Del 22 al 24 de Marzo	
Análisis de Reactivos y constitución de versiones definitivas de los exámenes.	Del 27 de Marzo al 10 de Abril	
Impresión y Distribución de exámenes, hojas de respuesta y materiales de apoyo.	Del 11 de Abril al 24 de Mayo	Realizada por CONALITEG
Periodo de Aplicación	Del 5 al 9 de Junio	
Lectura Óptica de las hojas de respuesta y escaneo	Julio	
Calificación	Agosto	
Publicación en Internet de resultados individuales y agregados por escuela.	1 de Septiembre	

Fuente: Tomado de *Evaluación Nacional de Logro Académico Informe técnico 2006*

El proceso de elaboración de la prueba ENLACE que se muestra en el Cuadro 21, duró prácticamente cinco meses y si consideramos que la aplicación de la prueba ENLACE 2007, inició el 23 de abril del 2007, sólo seis meses pasaron de la publicación de resultados de ENLACE 2006, a la aplicación de ENLACE 2007

Cabe destacar además que se le da mayor cantidad de tiempo a la impresión y distribución de exámenes, a la elaboración de hojas de respuesta y material de apoyo, que al proceso de jueceo y piloteo

Las etapas propuestas se centran en la selección del contenido para asegurar su representatividad y en la revisión técnica de los ítems con bases en las especificaciones elaboradas. Sin embargo no hay información o evidencia sobre su validez de contenido y de constructor de los reactivos elaborados.

Además es importante considerar que el Plan y Programa de Estudios vigente de educación básica a nivel primaria plantea como uno de sus primeros propósitos, que los niños

“adquieran y desarrollen las habilidades intelectuales (la lectura, y la escritura, la expresión oral, la búsqueda y selección de información, la aplicación de la matemática a la realidad), que les permita aprender permanentemente y con independencia, así como actuar con eficacia e iniciativa en las cuestiones prácticas de la vida cotidiana”, (SEP, 1993, p. 13)

Cabe revisar en qué medida las pruebas ENLACE, aplicadas hasta el momento consideran los objetivos planteados en el Plan y Programa de Estudio de la Educación Primaria. Por otra parte, el nuevo plan de estudios para la educación básica 2009, plantea como propósito central el desarrollo de competencias. Las competencias que se proponen desarrollar a través de todas las asignaturas son *Competencias para el aprendizaje permanente, competencias para el manejo de información, competencias para el manejo de situaciones, competencias para la convivencia y competencias para la vida en sociedad* (SEP, 2008b, p. 36-37). Ante lo cual surge una primera contradicción, si en los planes y programas de estudio de la educación primaria se establece el desarrollo de habilidades intelectuales y la aplicación de las matemáticas a la realidad, así como el desarrollo de competencias, ¿por qué ENLACE es una prueba centrada en el conocimiento y contenidos curriculares?

Una parte fundamental del proceso de elaboración y construcción de la prueba, también lo constituye la elaboración de la tabla de especificaciones a partir de la cual se define el contenido de los ítems. Conviene ahora ver, que tanto las tablas de especificaciones, como un elemento base en el proceso de elaboración de reactivos, contemplan los objetivos de los planes y programas de estudio.

### 3.1.7 Elaboración y selección de reactivos

La elaboración y selección de reactivos de la prueba ENLACE, se realiza con base en tres momentos

- 1) La definición del contenido a través de las tablas de especificaciones
- 2) Los procedimientos y normas para la elaboración de los reactivos
- 3) Jueceo de los reactivos mediante la incorporación de las preguntas elaboradas en un Banco de reactivos automatizado

#### 1) Definición del contenido y tabla de especificaciones

La elaboración de las matrices de los instrumentos, tablas de contenidos o de referencias, constituye una etapa metodológica importante en la elaboración de reactivos

La tabla de especificaciones, base para la elaboración de reactivos de la prueba ENLACE, fue elaborada en el 2006, por la Subsecretaría de Educación Básica (SEB), a través de la Dirección General de Desarrollo Curricular (DGDC), de acuerdo con el análisis realizado por el INEE en las pruebas EXCALE, lo que fue la base para definir los temas y contenidos principales de la prueba ENLACE

Las especificaciones tienen como propósito principal proporcionar a los redactores de ítems una indicación clara del contenido específico que debe ser evaluado. Esta matriz de especificaciones, es acorde con los contenidos establecidos en los planes y programas de estudios oficiales vigentes. El contenido de las pruebas responderá al diseño explicitado en la matriz de especificaciones para cada asignatura-grado.

Las tablas de contenidos o especificaciones, para la asignatura de matemáticas, está compuestas por tema, contenido y especificaciones respectivas para cada uno de los grados. Éstas son redactadas en términos de conductas medibles a través de pruebas objetivas.

La tabla de contenidos y especificaciones de Matemáticas para el 6° Grado de Primaria, contempla, 10 temas, cada uno de los cuales está dividido en diferentes contenidos, para los que se establecen especificaciones expresadas en términos de la tarea que los alumnos tienen que realizar

El Cuadro 22 muestra una síntesis de las especificaciones para la prueba de español 6° grado que contiene tipos de texto, contenidos curriculares, qué evaluar y especificaciones

**Cuadro 22**  
**Síntesis de las especificaciones**  
**Español 6° grado**

Tipos de texto	Contenidos curriculares	Qué evaluar	Especificaciones
<u>UU Textos Informativos</u> Artículo Informativo de reporte de investigación Noticia Entrada de diccionario Plano urbano Gráfica de barras	Artículo informativo  Estrategias de lectura  Pares de un texto Identificación de la estructura de los tipos de discurso	Comprensión global del texto Inferencias para la interpretación del significado global  Comprensión de la idea principal de un párrafo	Reconocer el significado de palabras o frases de uso poco frecuente o técnico, a partir de las pistas que da el texto para su dilucidación  Presentar en la base del reactivo un tecnicismo resaltado en negritas o cursivas, cuyo significado se pide dilucidar, a partir de la información incluida en el texto. No debe presentarse la palabra aislada
<u>Textos literarios narrativos, teatrales y poéticos</u> Relato cuento Relato leyenda Poema Guión de teatro	Interpretación y uso de palabras y frases a partir del significado global  Reconocimiento de las relaciones forma-significado entre las palabras.	Reconocimiento de la conclusión del reporte de investigación  Identificación del significado de las palabras o frases  Reconocimiento de prefijos para marcar oposición	Identificar el significado de una palabra del texto en la que esté presente un prefijo que exprese oposición o negación y que indique alguna característica, estado, contexto o peculiaridad en la investigación que se reporta
<u>Textos epistolares</u> Carta formal			
<u>Portadores</u> Cartel			

Fuente: elaborado con información de las Tablas de Contenidos Sexto Grado de Primaria

Con respecto a las especificaciones elaboradas para la prueba de Español 6° grado, podemos decir que los objetivos señalados en el Plan y Programa de Estudios de la Educación Primaria para el área de Español (SEP, 1993) establecen como una de sus prioridades, propiciar el desarrollo de las capacidades de comunicación de los niños. Sin embargo, en las especificaciones elaboradas difícilmente se advierte este objetivo. Además de que no se detectan especificaciones para obtener la reflexión sobre lo que se lee, que es otro de los aspectos importantes que también se expresan en dicho plan. Las especificaciones que sobre la lectura se dan todavía quedan en un nivel básico.

La distribución de reactivos para la prueba de Español 6° grado, con base en la tabla de especificaciones, quedó de la siguiente manera:

**Cuadro 23**

**Distribución de reactivos para la prueba de Español sexto grado**

**ENLACE BÁSICA ESPAÑOL**  
Distribución de Reactivos  
por Tipo de Texto y Línea de Evaluación  
Sexto de Primaria

Tipo de texto		Líneas de Evaluación			Total
		Funciones de la lectura, tipos de texto, características y portadores	Comprensión lectora	Reflexión sobre la lengua	
Textos Informativos	Reporte de investigación	1	4	1	6
	Noticia	1	3	2	6
	Entrada de diccionario	3	0	0	3
	Plano urbano	3	0	0	3
	Gráfica de barras	2	2	0	4
Textos Literarios: Narrativos, Teatrales y Poéticos	Cuento	5	6	0	11
	Leyenda	1	2	1	4
	Poema	3	0	1	4
	Guión de teatro	4	2	0	6
Textos Epistolares	Carta formal	1	7	1	9
Portadores	Cartel	2	2	4	8
<b>Total</b>		<b>26</b>	<b>28</b>	<b>10</b>	<b>64</b>

Fuente: SEP (s/f). *Evaluación Nacional del Logro Académico en Centros Escolares, ENLACE Educación básica*. Recuperado 7 de noviembre del 2009 de [http://enlace.sep.gob.mx/ba/viewimg.php?img=img/tablas/esp\\_p6.png](http://enlace.sep.gob.mx/ba/viewimg.php?img=img/tablas/esp_p6.png)

En esta distribución podemos observar que en materia de español, textos literarios y narrativos, en especial el cuento, los textos portadores de información, los reportes de investigación y las noticias, son los que ocupan el mayor número de reactivos

Habría que revisar que tan significativos resultan este tipo de textos para niños de 12 y 13 años de los diferentes contextos en que se aplica la prueba ENLACE

Así mismo, tenemos que el mayor número de reactivos es de comprensión lectora. De ahí el interés por analizar los reactivos de esta línea de evaluación

## **2) Procedimientos y normas para la elaboración de los reactivos**

Los reactivos para la prueba ENLACE son elaborados, de acuerdo a un "Manual de Normas" que presenta lineamientos según el campo de estudio que trata, el área o disciplina determinada y da estructura determinada, considerando los procedimientos para elaborar reactivos de opción múltiple

Cabe señalar que el CENEVAL, el INEE y la Dirección de la SEP, siguen la misma línea técnica para la elaboración de los reactivos, el procedimiento para las distintas instancias es el mismo y es CENEVAL quien proporciona la plataforma computacional para su elaboración y da los lineamientos y cursos para su elaboración

Para el caso de español, se distinguen dos tipos de reactivos el simple o independiente y el multirreactivo

El reactivo simple o independiente es el reactivo compuesto por la base constituida por una pregunta, afirmación, gráfico o enunciado que describe un problema o situación planteada explícitamente y las opciones de respuesta con cuatro opciones



El multirreactivo, permite evaluar de forma integrada una variedad de conocimientos. Para que un reactivo pueda ser considerado como multirreactivo, del planeamiento de su base (llamado reactivo padre), deben descomponerse como mínimo tres reactivos (denominados reactivos hijos).

Los lineamientos específicos para el área de español establecen que los textos informativos, deberán contener información fidedigna. Se debe evitar el uso de textos con temas polémicos, que utilicen un lenguaje discriminatorio y ofensivo en cuanto a género, raza, ciencia, etc.

Todos los reactivos de todas las pruebas, independientemente de la asignatura se elaboran con base en una plantilla. La plantilla para la elaboración de reactivos, presenta el siguiente formato<sup>60</sup>.

**TARJETA REACTIVO**

**CICLO ESCOLAR: 2008-2009**

**PROYECTO:**

<b>DESTINATARIO:</b>			
ID_1 TIPO_EXAMEN	ID_2 ASIGNATURA	ID_3 NIVEL:	ID_4 NP:
ID_5 RUBRO DEL TEMARIO	ID_6 ASPECTO DEL TEMARIO	ID_7 BLOQUE:	ID_8 TEMA ESPECÍFICO:
<p>REACTIVO</p> <p>RC:</p>			
ID_9 JUSTIFICACIÓN:			

<sup>60</sup> Cabe señalar que este formato es utilizado en todos los exámenes que se aplican a nivel nacional.

AUTOR	REFERENCIA BIBLIOGRÁFICA
FECHA	
ID_10 NUM REGISTRO	

Una vez concluida la elaboración de reactivos, pasan a la etapa de jueceo<sup>81</sup>

### 3) Jueceo de los reactivos

El jueceo de los reactivos de la prueba ENLACE se hace mediante la incorporación de las preguntas en un banco de reactivos automatizado conocido como BFAE el cual es proporcionado por el CENEVAL. Los jueces son los especialistas de la misma Dirección de Evaluación la SEP. Generalmente hay 7 jueces para la misma prueba. Dentro de esta plataforma hay opción para que los jueces emitan sus sugerencias y fundamentos.

Posteriormente se realiza un estudio piloto en una muestra de alumnos para valorar el comportamiento de los reactivos a fin de asegurar el cumplimiento de estándares de calidad internacional esperados. Para este piloteo se utilizan seis formas distintas de pruebas.

Con la información resultante en el jueceo y el piloteo, se define un marco metodológico y se realiza un análisis de los reactivos a través del *ítem análisis*.<sup>82</sup>

Con el *ítem análisis* se obtiene el porcentaje de alumnos que contestaron correctamente cada reactivo, el índice de discriminación y el *punto biserial* que se utiliza para saber si las personas "adecuadas" son las que obtienen las respuestas correctas, qué tanto poder predictivo tiene el reactivo y cómo puede contribuir a las predicciones (SEP, s/f).

<sup>81</sup> Se conoce como jueceo el jurado colegiado de los reactivos que conforman la prueba.

<sup>82</sup> El *ítem análisis* ITEMAN es un programa computacional que proporciona indicadores estadísticos clásicos sobre el funcionamiento de los ítems como la dificultad, el índice de discriminación, el *rbpls* (un coeficiente de correlación del ítem con el total de la prueba), y otros resúmenes estadísticos como los promedios y la proporción de personas que contestaron el test tanto por cada ítem como el total de la prueba. Cada indicador facilita referirse con mayores elementos a un ítem.

A continuación se muestran las tablas elaboradas por la SEP, con los índices de dificultad y discriminación aceptables para cada reactivo

Cuadro 24

**Índices de dificultad y discriminación aceptables para cada reactivo.**

**Índice de dificultad**

Ítems superados por el 85%	Muy fáciles
Ítems superados por el 50 al 85%	Relativamente fáciles
Ítems superados por el 15 al 50%	Relativamente difíciles
Ítems superados por el 0 al 15%	Difíciles o muy difíciles

Tomado de SEP (s/f) *Análisis de ítems Índice de dificultad Índice de discriminación* Versión en power point

**Índice de discriminación**

40 y más	Muy buenos ítems
30 a 39	Razonablemente buenos, pero sujetos a mejoramientos
20 a 29	Regulares deben mejorarse
Menos de 19	Deficientes deben descartarse o revisarse muy prolijamente

Tomado de SEP (s/f) *Análisis de ítems Índice de dificultad Índice de discriminación* Versión en power point

De esta manera, se seleccionan los mejores ítems a fin de integrar la prueba definitiva y contar con mayores elementos para fundamentar su inclusión en la versión final del examen

Esto es lo que se dice de la prueba ENLACE, sin embargo cabe destacar que además de los resultados obtenidos,<sup>83</sup> no encontramos información sobre aspectos pedagógicos implícitos en su elaboración y aplicación

<sup>83</sup> En la página de ENLACE <http://enlace.sep.gob.mx/ba/> se pueden consultar los resultados obtenidos, por alumno y por escuela desde el 2006 al 2009, siempre y cuando se tenga la clave correspondiente. Además se presentan las estadísticas que incluyen los resultados históricos nacionales y por entidad federativa, en medias y niveles de logro, una evolución de los resultados del 2006 al 2009, en donde se muestra el comportamiento estadístico de los alumnos en los diferentes niveles, desde el 2006 hasta el 2009. Sin embargo esta información está en términos estadísticos, de traducción para los docentes, alumnos y padres de familia. Las sugerencias que se hacen para el uso de resultados (SEP, 2006c) están avocadas a realizar recomendaciones y acciones generales, como el incrementar la capacidad de los padres para apoyar a sus hijos en las tareas educativas. Sin embargo no se encuentra información que de a conocer a los docentes y alumnos el por qué de los resultados obtenidos en términos de enseñanza y aprendizaje

Los elementos que caracterizan a la prueba ENLACE, junto con su proceso de elaboración y aplicación, permiten elaborar algunos cuestionamientos al respecto y poner de manifiesto sus concepciones implícitas

### 3.1.8 Intenciones explícitas, cuestionamientos que surgen y concepciones implícitas

Se dice que ENLACE es una prueba cuyo objetivo es obtener información adicional sobre el funcionamiento de la educación escolar en México a través de los aprendizajes en matemáticas y español de estudiantes que cursan el tercero, cuarto, quinto y sexto de primaria y tercero de secundaria de todo el país. Se concibe como un instrumento creado para promover la rendición de cuentas.

En la información disponible en algunos documentos<sup>84</sup> y en su sitio web,<sup>85</sup> se presentan las siguientes intenciones:

- Fortalecer el Sistema Nacional de Evaluación
- Mejorar los procesos educativos
- Difundir resultados a la sociedad, para mejorar calidad de la educación básica
- Proporcionar información diagnóstica del grado (nivel) en que los alumnos han adquirido los temas y contenidos de las asignaturas de español, matemáticas y ciencias

Las intenciones explícitas y su relación con las características que definen a la prueba ENLACE, permiten cuestionar la naturaleza de la prueba ENLACE como instrumento de evaluación para mejorar del proceso educativo al reflexionar sobre las siguientes cuestiones:

- Una prueba censal centrada en resultados cuantitativos del logro escolar, ¿puede dar información sobre el funcionamiento de la educación escolar y ayudar a mejorar el proceso educativo y de enseñanza en las aulas? ¿Es a través de evaluaciones estandarizadas a gran escala que se pretende fortalecer el sistema de evaluación? ¿De qué concepto de evaluación estamos hablando?, ¿Se puede llamar evaluación a un instrumento centrado en la medición matemática de resultados?

<sup>84</sup> Los documentos consultados fueron *Evaluación Nacional de Logro Académico en Centros Escolares* y el *Informe técnico 2006*. Documentos internos de trabajo en relación a la prueba 2006, proporcionados por la SEP en noviembre del 2006, pero que no son divulgados al público en general.

<sup>85</sup> El sitio web de ENLACE es <http://www.enlace.sep.gob.mx/>.

- ¿Logro académico y medición de resultados del trabajo escolar es igual a aprendizaje? ¿Cuál es el concepto de aprendizaje que subyace en la prueba ENLACE?
- Un diagnóstico de tipo factual, ¿de qué tipo de conocimiento habla? ¿El logro académico significa reportar la adquisición de conocimientos?
- ¿Qué es lo que permitirá mejorar la calidad de la educación, las pruebas estandarizadas, centradas en logros y resultados o la revisión de los procesos que dieron origen a esos logros y resultados?
- ¿Por qué el levantamiento de actas? ¿Se trata acaso de una cuestión confidencial o de una acción pedagógica?

Las características, intenciones y cuestionamientos a los que hemos hecho alusión permiten identificar algunas concepciones implícitas de esta prueba con respecto a la evaluación, al aprendizaje y a la mejora de la calidad

Si se considera a ENLACE como un instrumento para fortalecer un sistema de evaluación, tenemos que la evaluación a través de pruebas masivas estandarizadas, cuyos resultados se utilizan para premiar o castigar a los mejores,<sup>86</sup> más que una evaluación, es una manera de ajustar cuentas. En palabras de Bustamante y Caicedo

---

<sup>86</sup> La presidencia de México organiza una ceremonia especial para premiar a los mejores alumnos de la prueba ENLACE. De acuerdo a los resultados de la aplicación 2005-2006, la ceremonia se realizó el 15 de noviembre del 2007 <http://eh.ec.atl.presidencia.gob.mx/prensa/discursos/?contenido=32633&imprimr=true>. En esta ceremonia, según declaraciones de la Secretaría de Educación Pública, estuvieron 10 estudiantes de cada entidad federativa que obtuvieron los mejores resultados en la Prueba ENLACE 2007, de tercero a sexto de primaria y también de tercero de secundaria. Así como los 500 *mejores maestras y maestros* de escuelas públicas y privadas, en cuyas aulas se alcanzaron los mejores promedios en español y matemáticas, conjuntamente 92 de estas maestras y maestros son responsables de las aulas donde se obtuvieron las mejores calificaciones en ambas materias [http://www.sep.gob.mx/web/sep1/sep1\\_VersionJVM151107](http://www.sep.gob.mx/web/sep1/sep1_VersionJVM151107). La entonces titular de la SEP expresó: La calificación máxima que se puede alcanzar en ENLACE es de 936 puntos. Los niños que hoy lo acompañan, señor presidente, todos los que están aquí lograron el nivel de excelencia en la Prueba ENLACE. Son *los mejor calificados de todo el país* y estamos muy orgullosos. <http://www.presidencia.gob.mx/prensa/discursos/?contenido=28926> (las cursivas son nuestras). "Se premiará a las escuelas que hayan mejorado sus puntajes". *La Jornada* 9 de abril 2008 <http://www.vanguardia.com.mx/diario/noticia/politica/nacional/se-aplicara-la-prueba-enlace-en-escuelas-primarias-sep149121>. Al darse a conocer los resultados de la Prueba ENLACE 2009 en Educación Básica, el Secretario de Educación Pública, Maestro Alonso Lujambio, anunció que se otorgarán incentivos a los profesores cuyos alumnos hayan logrado los mejores promedios durante esta prueba. Para ello, dijo, se destinarán 900 millones de pesos, destinados por el Congreso de la Unión para 2009 (consultado en <http://www.sep.gob.mx/web/sep1/bolet2391009>, diciembre 2009). Estas acciones y los calificativos que en el se utilizan deja muy clara las intenciones e intereses a los que obedece la aplicación de esta prueba. Además queremos destacar los calificativos de "mejores maestros" o "mejores calificados" como la manera de referirse a quien obtuvo las más altas puntuaciones. Lo que también revela la manera en como se elaboran los juicios

---

(2005) "El paredón al que se enfrenta a los alumnos, a los maestros y a las escuelas, no se aleja en nada a un juicio público en el que el acusado debe defenderse" (p 110)

La caracterización de ENLACE como prueba "objetiva" hace concebir a la evaluación como un asunto técnico que justifica su cientificidad a través de la matematización. Lo que contribuye a promover un modelo empresarial de escuela, donde se enjuicia el comportamiento y se asignan premios<sup>87</sup> a partir de los resultados obtenidos

La estandarización que de ésta se hace, lleva a pensar que todos los niños independientemente de las características del contexto escolar, de las escuelas y de los maestros, reciben la misma enseñanza

La heterogeneidad de contextos en los que se aplica la prueba ENLACE, una prueba estandarizada, y homogénea, lleva a cuestionar la validez de sus resultados

Más que un diagnóstico del funcionamiento escolar, podemos considerar que la evaluación a través de pruebas a gran escala, lo que genera es una producción de una nueva manera de seleccionar y privilegiar conocimientos y aprendizajes, en función de lo que viene en el examen. Una nueva manera de juzgar al sistema educativo y sus actores, éste vale en la medida del resultado del examen. Una nueva forma de concebir y valorar la enseñanza, ésta vale en relación a su efectividad en preparar a los alumnos para el examen

Sin embargo, coincidimos con algunos autores en que la evaluación del aprendizaje no se puede dar a través de pruebas estandarizadas a gran escala. Cuando se evalúa el aprendizaje a través de pruebas, el principio que prevalece es una idea de aprendizaje reproductiva, basada en la memoria y en la repetición acrítica de la información (Dalmazo y Passos, 2001)

Dice Perrenoud (2008) las pruebas son concebidas con fines de clasificación, más que de identificación del nivel de conocimientos. Una prueba no dice mucho acerca de cómo

---

<sup>87</sup> Los alumnos con mejores calificaciones en la prueba ENLACE, recibieron además de un reconocimiento por parte del presidente de la República, una Enciclopedia Hispánica Platina, compuesta por 10 tomos. Ya partir de este ciclo escolar, los maestros cuyos alumnos mejoren los resultados en la prueba ENLACE, obtendrán un bono de bonificación

---

se lleva a cabo el aprendizaje y la construcción de conocimientos en cada alumno, sanciona los errores, sin darle los medios para comprenderlos y trabajarlos

Con base en lo anterior, consideramos que la evaluación va más allá de la examinación. La evaluación del rendimiento, más que un examen, debe constituirse para el educando una oportunidad para profundizar sobre qué, cómo y para qué aprende el estudiante. Se trata de una relación pedagógica que debe llevar a la confrontación, la interlocución y la actualización (Baudillo, 2008)

La mejora de la educación se dará con la mejora del aprendizaje y por ende de la enseñanza, pero la medición estandarizada centrada en resultados, no contribuye a esta función. Mejorar cualitativamente la enseñanza implica proporcionar datos, análisis e interpretaciones válidas y fiables que permitan forjarse una idea precisa del sistema educativo y sus componentes (Tiana, 1996, p. 46)

ENLACE, es una prueba enfocada al rendimiento y a la rendición de cuentas y no al aprendizaje ni a la mejora de la enseñanza, porque hasta el momento no se conocen acciones, derivadas de la aplicación de la prueba ENLACE que persigan este fin

De acuerdo con Meirieu (1992), un aprendizaje, se efectúa cuando un individuo recoge información de su entorno en función de un proyecto personal a través de la interacción y la utilización. Pruebas como ENLACE difícilmente conciben esta interacción

Si consideramos que aprender no es retener o adquirir, sino y comprender. La efectividad en el aprendizaje no se logra o promueve con la memorización y la repetición, sino con el desarrollo del pensamiento, la comprensión y la retroalimentación

Dice Bain (2006)

“Las personas aprenden más efectivamente cuando intentan resolver problemas intrigantes, atractivos o importantes, son capaces de hacerlos en entornos que los desafían, pero que les dan apoyo y en el que sienten que tienen el control sobre su propia educación. Pueden probar, fallar y recibir retroalimentación” (p. 124)

ENLACE, como estrategia para mejorar la calidad, viene a ser parte de una política mundial, que responde a las necesidades de un mundo globalizado, para modificar sistemas educativos. Lo que ha obligado a las autoridades educativas a una aplicación masiva de exámenes como justificación para evaluar la calidad.

Desde esta perspectiva en la era de los test, la calidad está asociada a la rendición de cuentas, con lo que se incita a los maestros a preparar para el examen, ahora que ha sido políticamente impuesto y burocráticamente aceptado Day (2006). Hoy en día, las instituciones "educativas" están más preocupadas por cumplir las expectativas de la eficiencia y eficacia, términos por los que hemos sido invadidos, que por el desarrollo de las capacidades humanas. La visión humanista de la educación ha cambiado y se ha suplantado por una visión pragmática y productivista. Consideramos que no es a través de exámenes masivos que miden, cuantifican y comparan que se va a mejorar la calidad, sino a través de un estudio de las necesidades y situaciones a resolver en materia educativa. Y aunque el concepto de calidad puede ser abordado desde múltiples perspectivas y resulta muy complejo tratar de definirlo, hablar de calidad, es hablar de una cualidad intrínseca a un proceso, más que alcanzar un resultado en relación a otros. Lo que lleva a pensar en un concepto de calidad como respuesta a las necesidades educativas.

Mejorar la calidad, dice Bustamante y Díaz (2003), "exige no sólo medir los resultados del logro escolar de los estudiantes sino recopilar información en torno a los procesos y recursos que contribuyen a generar dichos resultados" (p. 17).

Un examen como ENLACE, centrado en la medición de logros y no en el aprendizaje, lo que mide es adquisición, repetición y adiestramiento, no aprendizaje. La evaluación del aprendizaje a través de pruebas estandarizadas centradas en resultados viene a ser como decía Kilpatrick (1968): "Un modo científico, popular de abordar este proceso del aprender, ha sido tratar de reducir el pensar y el aprender a algo que es menos que pensar, tratar de colocar el pensar y el aprender en lo posible sobre una base mecanizada" (p. 37).

Un examen centrado en el aprendizaje y no en la mera adquisición de conocimiento debe reflejar si se quiere que los estudiantes recuerden, comprendan, apliquen, analicen, sintetizen o evalúen (Bain, 2006, p. 63).



Las concepciones implícitas que subyacen a las intenciones explícitas, llevan a pensar en que ENLACE más que un instrumento de evaluación que ayude mejorar los procesos educativos, viene a constituir un instrumento de control y juicio, que tecnifica la evaluación para premiar o castigar y legitimar un determinado tipo de conocimiento. En donde el conocimiento se concibe como mera adquisición y se deja fuera el proceso de construcción y retroalimentación que debe haber en todo proceso de evaluación del aprendizaje desde una visión formativa.

Hasta aquí tenemos que la llamada *Evaluación Nacional de Logro Académico en Centros Escolares*, no es más que una manera de justificar técnicamente una acción política centrada en los resultados y enfocada a la rendición de cuentas.

Un uso más formativo de la evaluación a gran escala, requiere de un análisis didáctico no sólo de sus resultados sino también de su proceso de elaboración, que es lo que no hemos encontrado hasta el momento. De ahí la necesidad de realizar un análisis didáctico del proceso de construcción de los reactivos que conforman la prueba.

### **3.2 Un análisis didáctico de los reactivos de comprensión lectora de la prueba ENLACE, 2008 para 6<sup>a</sup> grado**

#### **3.2.1 Generalidades de la prueba ENLACE 2008, para 6<sup>o</sup> grado**

La prueba de ENLACE para 6<sup>o</sup> grado 2008, se compone de 178 reactivos, de los cuales 81 son de español. De estos 81 reactivos de español, 63 son de comprensión lectora y 18 son de reflexión sobre la lengua.

Estos reactivos han sido clasificados de acuerdo con su grado de dificultad obtenido en el *Ítem análisis* y presentado en las *Sugerencias para el uso pedagógico de los resultados de ENLACE* para sexto de primaria 2008 (véase SEP, 2008).

Cuadro 25

**Clasificación de los reactivos de comprensión lectora de acuerdo a su grado de dificultad**

**Clasificación de los reactivos de comprensión lectora de acuerdo a su grado de dificultad.**

COMPRESIÓN LECTORA

1-D	2-A	3-D	4-C	5-A	7-C	8-A	10-C	11-D	12-C
33-B	34-A	35-D	36-B	37-D	38-D	40-C	41-C	42-B	43-B
44-D	45-D	46-D	47-A	48-D	49-D	50-D	51-D	52-A	53-A
102-A	103-D	104-C	110-C	111-B	112-C	116-B	117-C	118-C	142-C
143-A	144-A	145-B	147-C	148-A					

	Preguntas con grado de dificultad BAJO
	Preguntas con grado de dificultad MEDIO
	Preguntas con grado de dificultad ALTO

Fuente. SEP, 2008c

De acuerdo con la clasificación presentada en el Cuadro 25, tenemos que de los 45 reactivos de comprensión lectora

15 presentan un grado de dificultad alto

27 presentan un grado de dificultad medio

3 presentan un grado de dificultad bajo

Con base en el modelo propuesto en el capítulo dos, presentamos el análisis de la estructura didáctica de un 30% de los reactivos de comprensión lectora, los cuales se han seleccionado con base en el nivel de dificultad establecido

De tal manera que se presenta el análisis de

4 reactivos con grado de dificultad alta

8 reactivos con grado de dificultad media

1 reactivo con grado de dificultad baja

### 3.2.2 Análisis didáctico de algunos reactivos de comprensión lectora en la prueba ENLACE 2008

#### Reactivos con grado de dificultad alto

##### Pregunta 1

Lee el texto y contesta las tres preguntas siguientes.

En la ciudad de Chihuahua más de 30 toneladas de materiales de desecho industriales están siendo aprovechadas para la edificación de casas y guarderías infantiles funcionales y de bajo costo.

El Comité de Ingeniería Ambiental coordina los trabajos de construcción, usando materiales tales como tarimas de madera, cartón, plástico y papel.

Como parte de trabajo, la basura no peligrosa de la industria va formando paneles con tarimas de madera que son forradas en su interior y exterior con tela de algodón y rellenas con cartón de desperdicio que las hace térmicas. Después, se mezclan las paredes con materiales convencionales, como la arena y el cemento.

Hefeme, 10 de julio de 1.51

1. ¿Cuál es el tema central del texto?
- A) La utilización de basura no peligrosa para elaborar tarimas de madera, cartón, plástico y papel.
  - B) El aprovechamiento de los materiales industriales de desecho para la construcción de casas y guarderías.
  - C) El uso que se le da a la basura peligrosa y desechos industriales para edificar casas y guarderías.
  - D) El beneficio que se obtiene al producir tarimas de madera rellenas con cartón que las hace térmicas.

**Pregunta:** 1

**Grado de Dificultad:** Alto

**Tipo de Texto:** Reporte de investigación

**Título de la Lectura:** Aprovechamiento de desechos industriales

**Contenido Curricular:** Artículo Informativo (reporte de investigación); tema e ideas principales.

**Propósito:** Reconocer el tema global que trata un reporte de investigación.

**Respuesta Correcta:** B

Fuente: SEP, 2008c, p. 27

### Análisis de su estructura didáctica de la pregunta 1

<b>Tarea de evaluación</b>	Se trata de una tarea que implica comprensión pero no transferencia ni problematización, porque sólo exige establecer relaciones entre la información presentada, pero no plantea un problema a resolver ni requiere movilizar conocimientos, además de que no resulta significativa porque la situación que presenta es lejana a la realidad y vida de un niño de sexto grado
<b>Organización lógica del contenido</b>	El contenido es de tipo conceptual. Requiere dotar de significado a la información presentada y establecer interrelaciones entre los elementos básicos
<b>Organización metodológica</b>	Implica la creación de situaciones didácticas que recurran a estrategias de organización de la información, a la realización de inferencias y a la práctica de actividades simbólicas que permitan representar el contenido del texto
<b>Organización cognitiva</b>	Demanda el uso de procesos simples de ordenamiento de hechos y recuperación de la información relevante y de un proceso complejo de síntesis para dotar de significado a la información presentada
<b>Organización sociocultural</b>	Se ubica en un contexto extradisciplinar, porque se trata de cuestiones situadas dentro de un contexto real relacionadas con aspectos de su entorno, pero ajenas a la realidad de todos los estudiantes

### Variables de construcción del reactivo que afectan su resolución

<b>Objetivo del ítem</b>	El propósito del reactivo de "Reconocer el tema global de que trata un reporte de investigación establecido en las tablas de especificaciones" se cumple en la redacción de la pregunta, pues ésta es muy directa. ¿Cuál es el tema central del texto? Sin embargo al reparar en los distractores, vemos que sobre todo el de la respuesta correcta está redactado con las mismas palabras del texto. Lo que induce a pensar que más que el reactivo más que evaluar el reconocimiento del tema central del texto los distractores utilizados inducen a que lo que se evalúa es la capacidad de identificar relacionar las palabras utilizadas en el texto
<b>Expresiones socioculturales del lenguaje</b>	Se trata de una situación ubicada en un contexto real, el vocabulario utilizado y tema al que se alude, no es de manejo común para los niños de diferentes contextos, pues habría que pensar que si un niño de la sierra o del campo sabrá lo que significa desechos o materiales industriales, que de hecho es el término utilizado en el reactivo que tiene la respuesta correcta. Es por esto que el código sociolingüístico utilizado favorece desigualmente el éxito de los niños, ya que no tienen el mismo sentido para los niños de diferentes contextos socioculturales

<b>Características perceptuales y de redacción de los distractores</b>	El reactivo no tiene problemas perceptuales, si consideramos que el papel de un distractor es evitar dar con la respuesta correcta. El distractor que tiene la respuesta correcta contiene los mismos términos utilizados en el texto "materiales industriales", por lo que si el estudiante se da cuenta de esto es fácil que de con la respuesta correcta sin tener una comprensión global del texto que es el objetivo del ítem.
<b>Estrategias empleadas para su resolución</b>	Al considerar que la respuesta correcta contiene las mismas palabras del texto, es factible que los estudiantes que identifiquen esta característica acudan a la eliminación para dar con la respuesta correcta.
<b>Justificación de distractores</b>	De acuerdo con la información presentada en la <i>Sugerencias para el Uso pedagógico de los resultados de ENLACE</i> , la respuesta correcta es la <b>b</b> , pero ninguno de los demás distractores reporta un error en el manejo de los contenidos. Pues se trata sólo de diferentes maneras de nombrar el contenido del texto e incluyen un poco de más parafraseo.

### Pregunta 9

<p>Lee el texto y contesta las tres preguntas siguientes:</p> <p>El equipo de la Clínica Médica Nacional, Unidad La Raza, realizaron con éxito el día 21 de julio de 1988 el primer trasplante de corazón en el país y el primero en América Latina con ese resultado.</p> <p>El paciente que recibió el corazón sintió, Fernando Talpa, de 45 años de edad y quien había sido en la Ciudad de México un enfermo, estaba afectado severamente y condenado a morir en un plazo no mayor de 4 meses.</p> <p>El médico cirujano responsable de la operación ha cuidado con detalle la evolución de paciente y está satisfecho con los progresos mostrados por el mismo.</p>	
9.	<p>En la oración "...El médico cirujano responsable de la operación ha cuidado con detalle la evolución del paciente...", ¿qué acepción del diccionario es la adecuada para buscar el significado de la palabra resaltada con negritas?</p> <p>A) n. s. f. Cambio gradual, progresivo.</p> <p>B) fig. mudanza de conducta, de propósitos o de actitud.</p> <p>C) fig. desarrollo o transformación de las ideas o teorías.</p> <p>D) n. Pl. L. Conjunto de movimientos, ejercicios, maniobras de alguien o algo.</p>
<p><b>Pregunta:</b> 9  <b>Grado de Dificultad:</b> Alto  <b>Tipo de Texto:</b> Entrada de diccionario  <b>Título de la Lectura:</b> Trasplante de corazón  <b>Contenido Curricular:</b> Entrada de diccionario: selección de la acepción correcta a partir del contexto; utilizar la información (abreviaturas, códigos) que proporciona el diccionario.  <b>Propósito:</b> Identificar la acepción más adecuada para buscar el significado de una palabra de acuerdo al contexto en el que se maneja en un texto.  <b>Respuesta Correcta:</b> A  <b>Fuente:</b> SEP, 2008c, p. 32</p>	

### Análisis de la estructura didáctica de la pregunta 9

<b>Tarea de evaluación</b>	Se trata de una tarea reproductiva de ejercitación rutinaria, centrada en ejercicios que requieren la práctica de algún contenido escolar y que resulta poco significativas
<b>Organización lógica del contenido</b>	El contenido se refiere a un dato específico que requiere de un bajo nivel de abstracción
<b>Organización metodológica</b>	Requiere de situaciones didácticas que recurran a estrategias de repaso, de procedimientos de repetición simple y actividades simbólicas
<b>Organización cognitiva</b>	Implica el uso de procesos cognitivos simples de reconocimiento e identificación
<b>Organización sociocultural</b>	Está ubicado en un contexto disciplinar y cuestiones escolares

### Variables de construcción del reactivo que afectan su resolución

<b>Objetivo del ítem</b>	El objetivo del ítem de identificar la acepción más adecuada para buscar el significado de una palabra de acuerdo al contexto en el que se maneja en un texto, queda un poco confuso, porque aunque la pregunta está planteada en estos términos, para poderla contestar se requiere primero conocer el significado de evolución, por lo que no se sabe bien si lo que se quiere es evaluar el manejo del vocabulario o el manejo de acepciones en el diccionario
<b>Expresiones socioculturales del lenguaje</b>	Expresiones como acepciones y evolución, son palabras que resultan poco familiares y significativas para niños de contextos rurales lo que dificulta su comprensión y lo pone en situaciones de desigualdad
<b>Características perceptuales y de redacción de los distractores</b>	El reactivo no presenta problemas preceptuales ni de redacción, pero dos de las acepciones de los distractores son iguales, lo que confunde y ayuda a encontrar la respuesta correcta en lugar de impedirlo, pues si el alumno se da cuenta de esta característica, por lógica desecha estas dos opciones y puede llegar más fácilmente a encontrar la correcta
<b>Estrategias empleadas para su resolución</b>	Dadas las características de redacción de los distractores y la repetición de acepciones en dos de ellos, es fácil que los estudiantes den con la respuesta correcta por eliminación
<b>Justificación de distractores</b>	Si cada uno de los distractores tuviera una acepción diferente, los distractores ayudarían a reportar errores en el manejo del contenido, pero el que dos de los distractores tengan la misma acepción dificulta poder identificar dificultades en el manejo del contenido

### Pregunta 98

Lee con atención el poema y responde las siguientes cuatro preguntas.

#### La Enredadera

*Por el molino del huerto  
asciende una enredadera.  
El esqueleto de hierro  
va a tener un chal de seda.*

*Ahora verde, azul más tarde  
cuando llegue el mes de enero  
y se abran las campanillas  
como puñados de cielo.*

*Alma mía: ¿quién pudiera  
vestirse de enredadera!*

*Juana de Ibarbouro*

- 98 En el poema hay un verso resaltado en negritas. ¿a qué se refiere?
- A) A la reja del molino.
  - B) Al molino en una loma.
  - C) A las flores de la planta.
  - D) Al chal que está en el suelo.

Pregunta: 98

Grado de Dificultad: Alto

Tipo de Texto: Poema

Título de la Lectura: La enredadera

Contenido Curricular: Poema: versos, ritmo y rima, interpretación de recursos literarios (comparación y metáfora).

Propósito: Comprender el sentido de una metáfora en un poema.

Respuesta Correcta: A

SEP, 2008c, p. 50

### Análisis de la estructura didáctica de la pregunta 98

<b>Tarea de evaluación</b>	Se trata de una tarea que implica comprensión, porque implica establecer relaciones, pero no es problematizadora, no induce a movilizar conocimientos, ni implica una transferencia, por lo que no resulta significativa
<b>Organización lógica del contenido</b>	El contenido es de tipo conceptual, requiere establecer relaciones y dotar de significado a la información o material presentado para comprenderlo
<b>Organización metodológica</b>	Implica la creación de situaciones didácticas que recurran a estrategias de selección de información y al uso de procedimientos de parafraseo e inferencias y actividades de tipo simbólicas
<b>Organización cognitiva</b>	Demanda del proceso complejo de interpretación de textos
<b>Organización sociocultural</b>	Está ubicado en un contexto de tipo escolar pues se refiere a contenidos disciplinares relacionado con cuestiones que se realizan en la escuela

#### Variables de construcción del reactivo que afectan su resolución

<b>Objetivo del ítem</b>	Para que el propósito del ítem de "comprender el sentido de una metáfora en un poema," pueda cumplirse, requiere que se tengan conocimientos previos de lo que es una metáfora, de lo que es un verso y de lo que significa un poema
<b>Expresiones socioculturales del lenguaje</b>	Las expresiones usadas tanto en el poema como en el reactivo, adquieren sentido sólo para quien está familiarizado con enanos y enredaderas molinos huertos y enredaderas
<b>Características perceptuales y de redacción de los distractores</b>	El reactivo no presenta problemas perceptuales ni de redacción. Sólo que el último distractor, el D), resulta totalmente ilógico
<b>Estrategias empleadas para su resolución</b>	Si el estudiante no está familiarizado con el contexto al que se refiere el poema, ni tiene los conocimientos previos necesarios, puede ser que la estrategia a la que recurra sea la de adivinación
<b>Justificación de distractores</b>	Los distractores empleados no ayudan a identificar el error en el manejo de algún contenido



### Pregunta 147

Lee la carta y contesta a las siguientes cuatro preguntas

Toluca, Tlax., a 16 de febrero de 2005

Amable Prujandine-rosari

Sra. Olga Rios Vazquez  
PRESENTE

Lamento el suceso ocurrido mi hijo no tuvo intención de chocar contra su carro. Pongase en contacto conmigo, mi compañía de seguros se encargara del asunto y le pagará lo que corresponda. A mi me ocurrió algo parecido el mes pasado.

Le presento mis más sentidas disculpas por tan desafortunado acontecimiento. \_\_\_\_\_

Atentamente

F. Carlos Martínez Ramírez

147. En la carta hay una línea al final, ¿cuál es la despedida adecuada para completar esa carta?

A) Hasta pronto. Oquita  
B) Hasta luego y gusto en saludarte.  
C) Sin otro particular, reciba un cordial saludo.  
D) Con cariño me despido, enviándole todo mi afecto.

**Pregunta:** 147  
**Grado de Dificultad:** Alto  
**Tipo de Texto:** Carta formal  
**Título de la Lectura:** Carta para la Sra. Olga Rios Vazquez  
**Contenido Curricular:** Identificación de la estructura de los tipos de discurso  
**Propósito:** Distinguir las características distintivas de la carta formal.  
**Respuesta Correcta:** C

Fuente: SEP, 2008c, p. 68

### Análisis de la estructura didáctica de la pregunta 147

<b>Tarea de evaluación</b>	Se trata de una tarea reproductiva de ejercitación rutinaria que resulta poco o nada significativa, porque solicita la reproducción de saberes aislados y está centrada en ejercicios que requieren de la práctica.
<b>Organización lógica del contenido</b>	Presenta un contenido centrado en la adquisición de hechos específicos que requiere de un bajo nivel de abstracción.
<b>Organización metodológica</b>	Este contenido factual, requiere de la creación de situaciones didácticas que recurran a estrategias de repaso y recirculación de la información, a procedimientos de repetición y acumulación de la información para distinguir las características distintivas de la carta formal y actividades reproductivas.

<b>Organización cognitiva</b>	Implica el uso de procedimientos simples de reconocimiento e identificación
<b>Organización sociocultural</b>	Se ubica en un contexto disciplinar de tipo escolar

**Variables de construcción del reactivo que afectan su resolución**

<b>Objetivo del ítem</b>	El objetivo del ítem de "distinguir las características distintivas de la carta formal", queda un poco confuso y al aire porque, en el reactivo no se menciona nada de carta formal. Se pide que se identifique o reconozca la despedida adecuada, pero faltaría explicar ¿qué se entiende por adecuada? Lo cual además resulta un tanto subjetivo, porque no hay un punto de referencia para definir lo adecuado, y lo que para un niño puede ser adecuado, quizá no coincida con lo que el elaborador del reactivo, está pensando como adecuado. Por lo tanto más que distinguir las características distintivas de una carta formal lo que se está evaluando es el manejo de cierto vocabulario.
<b>Expresiones socioculturales del lenguaje</b>	Las expresiones de lenguaje utilizadas demandan el manejo y comprensión de cierto vocabulario, porque palabras como <i>indemnización</i> no resultan familiares para todos los contextos socioculturales. Por lo que las preguntas que se formulan, no tienen el mismo sentido para las distintas clases sociales. Además, el contexto en el que se ubican y la situación a la que se refiere es lejana a la vida de un niño de sexto grado.
<b>Características perceptuales y de redacción de los distractores</b>	El reactivo no presenta problemas perceptuales de los distractores, pero todos los distractores se refieren a lo mismo, todos aluden a una despedida.
<b>Estrategias empleadas para su resolución</b>	Si el alumno alcanza a deducir lo que sería una despedida "adecuada" para el que elaboró el reactivo, lo más probable es que acuda a la adivinación como estrategia, porque todos los distractores se refieren a despedidas, faltaría sólo adivinar cuál es la adecuada.
<b>Justificación de distractores</b>	Los distractores empleados no permiten identificar el error en algún manejo del contenido, si acaso en la falta de manejo de cierto vocabulario.

Reactivos con grado de dificultad medio

Pregunta 39

La siguiente información se presenta en una gráfica con los datos, lee la información y contesta las siguientes cuatro preguntas.

Alumnos del 6° "A"

Categoría	Niños 10 años	Niños 11 años
Grupos	8	2
Boys	4	10

La gráfica anterior muestra la distribución de los alumnos que se inscribieron al grupo 6° "A" en la escuela primaria "Miguel Alemán" en la zona sur de la ciudad y sexo. Como puede observarse: 1) 8 son niñas y 4 son niños; 2) 52% del total de alumnos son niños que se cumplen 11 años y 48% son niñas que no han cumplido los 11 años; 3) 20% son niñas que no han cumplido los 10 años y el 20% son niños que ya se cumplen.

39. ¿Cuántos alumnos se inscribieron en el grupo 6° "A"?

A) 2                      D) 8                      C) 10                      D) 25

**Pregunta:** 39  
**Grado de Dificultad:** Medio  
**Tipo de Texto:** Gráfica de barras  
**Título de la Lectura:** Alumnos del 6° "A"  
**Contenido Curricular:** Gráfica de barras: búsqueda y localización de información específica.  
**Propósito:** Interpretar la información representada en una gráfica de barras con base en un texto breve sobre porcentajes o cifras.  
**Respuesta Correcta:** D

Fuente: SEP, 2008c, p. 41

Análisis de la estructura didáctica de la pregunta 39

<b>Tarea de evaluación</b>	Se trata de una tarea problematizadora que implica una comprensión, pero no una transferencia. Puede resultar significativa, pero no moviliza conocimientos previos.
<b>Organización lógica del contenido</b>	Presenta un contenido procedimental, que requiere de una acción sistematizada.
<b>Organización metodológica</b>	Implica una organización metodológica para la creación de situaciones didácticas que recurra al empleo de estrategias de construcción y procedimientos de elaboración. Así como de la práctica de actividades relacionadas con personas.
<b>Organización cognitiva</b>	Demanda el uso de procesos simples de ordenamiento, descripción y recuperación de la información.
<b>Organización sociocultural</b>	Se ubica en un contexto escolar y disciplinar.



<p>43 ¿En cuál de las siguientes opciones se resume el inicio y el final del cuento?</p> <p>A) En la plaza del pueblo había un árbol al que le pusieron una cerca alrededor, con un cartel. Después, él veía con sus ramas al oír los secretos de los niños.</p> <p>B) En la plaza del pueblo había un precioso, pero viejo árbol, que estaba triste porque los niños ya no se sentaban a él. Después de que unos niños se subieron, hacía magia y se convirtió en un árbol que conocía de todos.</p> <p>C) Guillermo le contó a Agustín que él quería ganar muchas veces en el juego de las canicas para que Jorge no se riera de él. Después de ganar, Carlota se hizo su amiga.</p> <p>D) Guillermo estaba muy contento porque Carlota le regaló una canica preciosa que él guardó como "la canica de la buena suerte". Después, le contó a Agustín lo que pasó.</p>
<p><b>Pregunta:</b> 43  <b>Grado de Dificultad:</b> Medio  <b>Tipo de Texto:</b> Cuento  <b>Título de la Lectura:</b> El árbol mágico  <b>Contenido Curricular:</b> Cuento y leyenda; apreciación literaria; título, personajes, inicio, desarrollo, final.  <b>Propósito:</b> Identificar la oración que sintetice el planteamiento inicial y el desenlace de un cuento.  <b>Respuesta Correcta:</b> B</p> <p>Fuente: SEP, 2008c, p. 44</p>

#### Análisis de la estructura didáctica de la pregunta 43

<b>Tarea de evaluación</b>	Se trata de una tarea que implica una comprensión porque requiere establecer relaciones entre la información presentada pero no es problematizadora ni requiere de la transferencia o de la movilización de conocimientos previos.
<b>Organización lógica del contenido</b>	Presenta un contenido conceptual que requiere dotar de significado a la información presentada para comprenderla.
<b>Organización metodológica</b>	Requiere de una organización metodológica para la creación de situaciones didácticas donde se empleen estrategias de organización de la información y procedimientos de parafraseo.
<b>Organización cognitiva</b>	Demanda el uso de procedimientos simples de ordenamiento y recuperación de la información, así como de un proceso complejo de síntesis.
<b>Organización sociocultural</b>	Se ubica en un contexto escolar y extradisciplinar por ser el tipo de lecturas que se trabajan en la escuela.

**Variables de construcción del reactivo que afectan su resolución**

<b>Objetivo del ítem</b>	El objetivo del ítem de identificar la oración que sintetice el planteamiento inicial y desenlace de un cuento, va de acuerdo con la pregunta que se formula
<b>Expresiones socioculturales del lenguaje</b>	El reactivo no presenta problemas de manifestaciones lingüísticas de difícil comprensión para los diferentes contextos
<b>Características perceptuales y de redacción de los distractores</b>	No se perciben problemas perceptuales ni de redacción en los distractores
<b>Estrategias empleadas para su resolución</b>	Si el alumno no alcanza a comprender el significado y sentido del texto lo más probable es que acuda a la adivinación como estrategia para resolver el ítem
<b>Justificación de distractores</b>	Los distractores que no contienen la respuesta correcta lo único que pueden denotar es falta de comprensión del sentido del texto

**Pregunta 103**

Un maestro llevó a sus alumnos un cartel les pidió que investigaran, con la misma información para difundirla entre sus intereses en asistir a un sitio. Cesava con atención el cartel y responde las siguientes ocho preguntas.

La Secretaría de Educación Pública, en colaboración con el Consejo Nacional para la Cultura y las Artes te invita a divertirla con tu familia el próximo

Jueves 31 de agosto de 2006

Visita 17:00 a 21:00 hrs.

Mariposa, Museo Infantil, Bosque de Chapultepec



**MARIPOSA  
MUSEO INFANTIL**

IMAX  
MEGAPANTALLA

Arrecife de Corales 18:30 hrs.  
3D Manía II 19:30 hrs.

DOMO  
DIGITAL

Vida en otros Mundos 20:30 hrs.

Además cada uno de los integrantes de tu familia podrá participar en las rifas que se llevarán a cabo durante la visita

Recoge tus voleros en la dirección de tu escuela

**CUPO LIMITADO**

Pregunta por los paquetes especiales  
para alumnos con  
Promedio de 10

Porque un país con cultura es la diferencia



CONACULTA

**Pregunta 103**

<p><b>103</b> ¿A quién está dirigido el cartel?</p> <p>A) A las niñas y niños con promedio de 10.          B) A los profesores de los niños aplicados.          C) A las familias de los niños y niñas.          D) A los niños, niñas y sus familias.</p>
<p><b>Pregunta:</b> 103  <b>Grado de Dificultad:</b> Medio  <b>Tipo de Texto:</b> Cartel  <b>Título de la Lectura:</b> Mariposa museo infantil  <b>Contenido Curricular:</b> Canal: descripción y persuasión; emisor, mensaje principal, condiciones y vigencia.  <b>Propósito:</b> Identificar al público potencial de un cartel a partir del mensaje que emite.  <b>Respuesta Correcta:</b> D</p> <p>Fuente SEP, 2008c, p 54</p>

**Análisis de la estructura didáctica de la pregunta 103**

<b>Tarea de evaluación</b>	Se trata de una tarea que requiere de una comprensión, pero no es problematizadora, ni implica una transferencia, ni el uso de conocimientos previos
<b>Organización lógica del contenido</b>	Presenta un contenido conceptual que requiere establecer relaciones entre la información que se presenta para darle significado
<b>Organización metodológica</b>	Requiere de situaciones didácticas con una organización metodológica donde se apliquen estrategias de selección y clasificación de la información y procedimientos de identificación de palabras claves y categorización de la información y actividades de tipo práctico
<b>Organización cognitiva</b>	Implica el uso de procesos simples de reconocimiento, identificación, ordenamiento y recuperación de la información
<b>Organización sociocultural</b>	El contexto en el que se ubica es extradisciplinar, vinculado a situaciones reales, pero que no resultan familiares o cercanas a todos los contextos



**Variables de construcción del reactivo que afectan su resolución**

<b>Objetivo del ítem</b>	El objetivo del ítem de "identificar al público potencial de un cartel a partir del mensaje que emite", no va de acuerdo con la pregunta que se realiza, porque ésta requiere inferir la respuesta, ya que no está especificado en el cartel
<b>Expresiones socioculturales del lenguaje</b>	En el cartel existen ciertas expresiones como IMAX, MEGAPANTALLA, DOMO que no tienen el mismo sentido para los distintos contextos socioculturales. Quizá podría haberse utilizado un cartel con un mensaje más común a todos los contextos socioculturales
<b>Características perceptuales y de redacción de los distractores</b>	El distractor C y D, prácticamente tienen la misma información, por lo que confunde y no denota un error en el manejo del contenido, el A y B, prácticamente no pueden ser porque esta información no existe en el cartel
<b>Estrategias empleadas para su resolución</b>	Si el alumno se percata que el distractor A y B, son prácticamente ilógicos y el C y D se parecen, es muy probable que recurra a estrategias de eliminación o adivinación
<b>Justificación de distractores</b>	Los distractores empleados, tal como están redactados no denotan error en el manejo de contenido

**Pregunta 104**

<p>104. ¿Qué información utiliza el cartel para convencer al público para asistir al evento?</p> <p>A) Que se proyectará <i>Vida en otros mundos</i> en el Domo digital.</p> <p>B) Que puedes ir por tus entradas a la escuela.</p> <p>C) Que habrá rifas y también paquetes especiales</p> <p>D) Que son cinco horas continuas de actividades</p>
<p><b>Pregunta: 104</b>  <b>Grado de Dificultad: Medio</b>  <b>Tipo de Texto: Cartel</b>  <b>Título de la Lectura: Mariposa museo infantil</b>  <b>Contenido Curricular: Identificación del propósito de la lectura y el texto.</b>  <b>Propósito: Reconocer el elemento persuasivo de un cartel publicitario.</b>  <b>Respuesta Correcta: C</b></p> <p>Fuente SEP, 2008c, p 55</p>

#### Análisis de la estructura didáctica de la pregunta 104

<b>Tarea de evaluación</b>	Se trata de una tarea que implica una comprensión porque exige establecer nexos
<b>Organización lógica del contenido</b>	Presenta un contenido conceptual que requiere establecer relaciones entre la información que se presenta para dotarla de significado
<b>Organización metodológica</b>	Requiere de situaciones didácticas donde se utilicen estrategias didácticas de organización de la información y procedimientos de identificación de palabras claves y categorización, con actividades simbólicas y prácticas
<b>Organización cognitiva</b>	Implica el uso de procesos simples de reconocimiento, identificación, ordenamiento y recuperación de la información
<b>Organización sociocultural</b>	El contexto en el que se ubica es extradisciplinar, vinculado a situaciones reales, pero que no resultan familiares o cercanas a todos los contextos

#### Variables de construcción del reactivo que afectan su resolución

<b>Objetivo del ítem</b>	El objetivo del ítem de "reconocer el elemento persuasivo de un cartel publicitario" va de acuerdo con la pregunta, siempre y cuando el alumno comprenda lo que es un elemento persuasivo
<b>Expresiones socioculturales del lenguaje</b>	En el cartel existen ciertas expresiones como IMAX, MEGAPANTALLA, DOMO que no tienen el mismo sentido para los distintos contextos socioculturales. Quizá podría haberse utilizado un cartel con un mensaje más común a todos los contextos socioculturales
<b>Características perceptuales y de redacción de los distractores</b>	Los distractores no presentan problemas perceptuales ni de redacción
<b>Estrategias empleadas para su resolución</b>	Dado que en tres de los distractores sólo una palabra está bien escrita y las otras dos son incorrectas, es probable que el alumno recurra a la eliminación para dar con la respuesta correcta
<b>Justificación de distractores</b>	Los distractores utilizados revelan es dificultad para localizar la información que se pide

Pregunta 110

EL UNIVERSAL

El Universal / Ciudad de México, Lunes 21 de agosto de 2006. Cuauhtémoc Blanco recibió su alta médica y a partir de este martes pasará a menos del preparador físico Axel Derbaum, mientras que el viernes será evaluado para definir si completa las acciones de fútbol con el equipo Águilas del América.

El doctor Alfonso Díaz declaró en entrevista que "médicamente a partir de mañana pasa con el preparador físico, con lo cual se le da su alta médica".

Blanco fue operado de la rodilla derecha y está en proceso de rehabilitación para integrarse al equipo que lleva una derrota, un empate y un triunfo en lo que se del Torneo Apertura 2006 de Fútbol Mexicano.

"Nosotros dijimos que en un promedio de cinco semanas estaría saliendo de la parte médica y esas dos semanas esas quinta semana", refirió el médico para indicar que se está cumpliendo con los tiempos de recuperación del jugador.

Respecto a la eventualidad de que Cuauhtémoc Blanco vuelva a sentir esta lesión, respondió que "médicamente existe el riesgo normal para una persona que juega al fútbol".

El médico de la Águilas del América aseguró que el jugador no siente agura molesta en su rodilla operada, y prueba de ello, dijo, es que ha realizado pequeños partidos de fútbol y o ha hecho bien.

110. Elige la opción que contiene la narración del hecho que se trata en la noticia.
- A) Cuauhtémoc Blanco será operado nuevamente y estará en proceso de rehabilitación, ya que los últimos días ha presentado muchas molestias en la rodilla derecha.
  - B) Cuauhtémoc Blanco jugará con su equipo en el próximo partido y en el Torneo Apertura 2006, ya que se encuentra recuperado de la lesión que sufrió en la rodilla izquierda.
  - C) Cuauhtémoc Blanco fue operado de la rodilla derecha; el martes será dado de alta y lo evaluará su preparador físico para ver si ya puede integrarse nuevamente a su equipo.
  - D) Cuauhtémoc Blanco será dado de alta para tomarse unas vacaciones y después integrarse al equipo que lleva una derrota, un empate y un triunfo en lo que va del Torneo Apertura 2006.

Pregunta: 110

Grado de Dificultad: Medio

Tipo de Texto: Noticia

Título de la Lectura: Cuauhtémoc Blanco se recupera y pronto volverá a la cancha

Contenido Curricular: Noticia; Ideas principales o puntos de vista; lugar, tiempo y participantes.

Propósito: Reconocer los hechos que se reportan y que responden a la pregunta ¿qué sucedió?, en una noticia

Respuesta Correcta: C

Fuente: SEP, 2008c, p. 58

Análisis de la estructura didáctica de la pregunta 110

<b>Tarea de evaluación</b>	Se trata de una tarea que implica la comprensión, porque requiere establecer relaciones entre la información
<b>Organización lógica del contenido</b>	Presenta un contenido conceptual que requiere dotar de significado al material presentado
<b>Organización metodológica</b>	Requiere la creación de situaciones didácticas donde se empleen estrategias de selección y organización de la información y procedimientos donde se identifiquen hechos e ideas, así como la realización de actividades simbólicas y prácticas

<b>Organización cognitiva</b>	Implica el uso de procedimientos simples de reconocimiento, identificación, organización y recuperación de la información
<b>Organización sociocultural</b>	El contexto en el que se ubica es extradisciplinar, vinculado a situaciones reales

**Variables de construcción del reactivo que afectan su resolución**

<b>Objetivo del ítem</b>	El objetivo del ítem de reconocer los hechos que se reportan y que respondan a la pregunta ¿qué sucedió? en una noticia, va de acuerdo con la pregunta formulada
<b>Expresiones socioculturales del lenguaje</b>	La situación y lenguaje utilizado como <i>rehabilitación</i> , y <i>preparador físico</i> , no son situaciones y expresiones fácilmente comprensibles para todos los contextos socioculturales, podrían haberse empleado expresiones más comunes como entrenador, o mejora. Así como noticias más de uso común, que no favorezcan de manera desigual el éxito de los niños de diferente sector sociocultural
<b>Características perceptuales y de redacción de los distractores</b>	Los distractores están elaborados, prácticamente con la misma información, sólo que acomodada de diferente manera y con algunos hechos de más o de menos. El distractor que contiene la respuesta correcta, está redactado casi de la misma manera que el texto, por lo que es fácil dar con éste
<b>Estrategias empleadas para su resolución</b>	Si el alumno se percató que el distractor que tiene la respuesta correcta está redactado casi de igual manera que en el texto, y que los demás distractores, contienen hechos que no vienen en el texto es probable que recurran a la eliminación como estrategia para dar con la respuesta correcta
<b>Justificación de distractores</b>	Los distractores no permiten identificar error en el manejo del contenido, porque presentan casi la misma información, pero acomodada de diferente manera y con algunos hechos de más o de menos

**Pregunta 116**

Lee el texto y contesta las siguientes preguntas.

**El Toro de Creta**  
(fragmento)

El Rey Minos de Creta parece é un día creóse resacaudo a Nestoro, Dios del Mar.

Creó el monstruo a Creta los bueyes más robustos de la Tierra, pero fingióle la muerte a Dios del Dios.

Por eso un día que el mar se desbordó y que le masaron a Nestoro y a sus hijos en las olas que se quebraron en las rocas del mar. El Dios del Mar que era un monstruo tan magnífico como el que se lo creó.

En las fiestas, un individuo llamado Minos que el Rey Minos le dio el nombre, ordenó se matara a un buey.

Incapaz de la ordenar, Minos le dio el nombre a un buey y lo dio a él el nombre de un buey.

Euristo, rey de Creta, encargó a Minos que le traía a Creta y lo traía a su palacio al lado del mar y dijo que era más difícil que matarlo.

Minos le dio el nombre a un buey y lo dio a él el nombre de un buey y lo dio a él el nombre de un buey.

Creó el monstruo que se lo creó en una gran ciudad llamada Minos y lo dio a él el nombre de un buey y lo dio a él el nombre de un buey.

Layna y Orosía, Luis Manuel  
Academia de Estudios de Historia y Geografía  
Instituto Tecnológico y de Estudios Superiores de Occidente

\* Debe leerse con atención, caso por caso.

---

116. ¿En cual de las siguientes opciones se menciona un hecho fantástico ocurrido en el relato?

A) "...hizo traer a Creta los bueyes más robustos de la Tierra..."

B) "...hizo surgir del mar a un magnífico toro..."

C) "...en la ceremonia ritual ordenó se matara a otro buey..."

D) "...llegaron sanos y salvos al palacio de Euristo..."

---

**Pregunta:** 116  
**Grado de Dificultad:** Medio  
**Tipo de Texto:** Leyenda  
**Título de la Lectura:** El toro de Creta  
**Contenido Curricular:** Cuento y leyenda; apreciación literaria, título, personajes, inicio, desarrollo, final y enseñanza.  
**Propósito:** Elegir la opción que exprese, en un texto, el origen de un fenómeno natural o extraño mediante elementos fantásticos.  
**Respuesta Correcta:** B

Fuente SEP, 2008c, p 63

**Análisis de la estructura didáctica de la pregunta 116**

<b>Tarea de evaluación</b>	Se trata de una tarea que implica la comprensión de lo que es un hecho fantástico y el establecimiento de relaciones en la información presentada. Pero no es problematizadora ni lleva a la transferencia.
<b>Organización lógica del contenido</b>	Presenta un contenido conceptual que requiere dotar de significado a la información presentada.
<b>Organización metodológica</b>	Requiere de la creación de situaciones didácticas con estrategias de clasificación de la información, mediante procedimientos de categorización de hechos y actividades simbólicas.
<b>Organización cognitiva</b>	Demanda de procesos simples de reconocimiento, identificación y recuperación de la información.
<b>Organización sociocultural</b>	Está ubicado en un contexto disciplinar de tipo escolar, porque está relacionado con contenidos escolares.

**Variables de construcción del reactivo que afectan su resolución**

<b>Objetivo del ítem</b>	El objetivo del ítem de elegir la opción que exprese en un texto el origen del fenómeno natural o extraño mediante elementos fantásticos
<b>Expresiones socioculturales del lenguaje</b>	No se observan problemas de lenguaje
<b>Características perceptuales y de redacción de los distractores</b>	Los distractores no presentan problemas perceptuales ni de redacción
<b>Estrategias empleadas para su resolución</b>	Lo que el alumno tiene que hacer es mero reconocimiento de un hecho fantástico. Como todos los distractores están elaborados con información del texto, es probable que si el alumno no reconoce lo que es un hecho fantástico trate de llegar a la respuesta correcta por adivinación, pero si reconoce lo que es un hecho fantástico, lo hará por eliminación descartando a los que no representen esta característica.
<b>Justificación de distractores</b>	Los distractores utilizados, lo más que representan es la ausencia del manejo de lo que es un hecho fantástico. No ayudan a identificar algún otro error en el manejo del contenido.

**Pregunta 142**

Lee la carta y contesta las siguientes cinco preguntas.

Ciudad, México, a 13 de agosto de 2006

1. **Destinatario:** Alma Beltrán Ortegón  
Directora de la Escuela Primaria "Fresca y Limpia"

2. **Remitente:** Estimada mi mamá

3. **Contenido:** A nombre del estudiante del grupo "A" por un año de haberme en la escuela, me solicito su intervención para la visita al banco educativo "Nuevo Día", con la intención de realizar una actividad que porge aprenda los contenidos adquiridos en el aula, además de participar en actividades de recreación y juegos que desarrollen mis habilidades y capacidades.

Solicito que por favor me envíe una autorización de permiso para poder salir a realizar esta actividad y poder participar en ella.

Le agradezco la atención prestada esperando tener una respuesta positiva.

4. **Asunto:** Atentamente  
Florencia Rodríguez Alvarado

142. En la carta, donde dice "...solicito su intervención...", ¿a que parte de la carta se refiere?  
A) Destinatario.  
B) Remitente.  
C) Petición.  
D) Saludo.

**Pregunta:** 142  
**Grado de Dificultad:** Medio  
**Tipo de Texto:** Carta formal  
**Título de la Lectura:** Carta para la Profra. Alma Estrada Ortega  
**Contenido Curricular:** Carta formal: fecha, destinatario, saludo, propósito, desarrollo y final; del sobre: datos del destinatario y del remitente.  
**Propósito:** Reconocer los elementos estructurales de la carta formal.  
**Respuesta Correcta:** C

Fuente: SEP, 2008c, p. 65

#### Análisis de la estructura didáctica de la pregunta 142

<b>Tarea de evaluación</b>	Se trata de una tarea reproductiva de ejercitación rutinaria y poco significativa, centrada en ejercicios que requieren la práctica de la identificación de los elementos estructurales de una carta. Situación que resulta un tanto ajena a la realidad de los niños, pues generalmente ellos no ocupan este medio de comunicación para dirigirse a las autoridades de su escuela o pedir algún permiso. Además de que pocas veces son los encargados de pedir estas autorizaciones.
<b>Organización lógica del contenido</b>	Presenta un contenido centrado en hechos, que requiere de un bajo nivel de abstracción, pues solamente tiene que identificar características.
<b>Organización metodológica</b>	Requiere de la creación de situaciones didácticas que recurran a estrategias de repaso y rearticulación de la información, con procedimientos de repetición y actividades de tipo reproductivas, sobre todo en cuestiones de vocabulario.
<b>Organización cognitiva</b>	Implica el uso de procesos cognitivos simples de observación, reconocimiento e identificación.
<b>Organización sociocultural</b>	Se ubica en un contexto de tipo escolar, relacionado con contenidos disciplinares.

#### Variables de construcción del reactivo que afectan su resolución



<b>Objetivo del ítem</b>	El objetivo del ítem de reconocer los elementos estructurales de la carta formal, va de acuerdo con la pregunta formulada, pero se requiere que el alumno maneje cierto vocabulario. Por lo que más que evaluar el reconocimiento de las partes de una carta, lo que se está evaluando es el manejo de vocabulario.
<b>Expresiones socioculturales del lenguaje</b>	El reactivo no presenta problemas de lenguaje. Se utiliza un lenguaje un tanto rebuscado y de uso poco frecuente como el término de <i>intervención, respuesta positiva</i> . No es una carta de uso muy común en todos los contextos socioculturales. Por lo tanto, este tipo de reactivos, no tiene el mismo sentido para los alumnos de los diferentes contextos socioculturales.



<b>Características perceptuales y de redacción de los distractores</b>	El reactivo prácticamente no presenta problemas perceptuales ni de redacción de los distractores
<b>Estrategias empleadas para su resolución</b>	Si el alumno no maneja el vocabulario empleado, lo más probable es que recurra a la adivinación como estrategia para la resolución de este reactivo
<b>Justificación de distractores</b>	Los distractores empleados. Lo que pueden revelar es dificultad en la identificación de los nombres y significado de los elementos de una carta

### Reactivos con grado de dificultad bajo

#### Pregunta 10

10. ¿Cuál de los siguientes símbolos representa en el mapa la ubicación de un restaurante?

A)  B) 

C)  D) 

**Pregunta: 10**  
**Grado de Dificultad: Bajo**  
**Tipo de Texto: Plano urbano**  
**Título de la Lectura: Plano para llegar a la casa de Jorge**  
**Contenido Curricular: Plano urbano: búsqueda y localización de información específica.**  
**Propósito: Identificar el significado de una acotación en un plano.**  
**Respuesta Correcta: C.**

Fuente: SEP, 2008c, p 34

#### Análisis de la estructura didáctica de la pregunta 10

<b>Tarea de evaluación</b>	Se trata de una tarea que implica una comprensión y transferencia, porque exige establecer nexos y aplicar un saber, aunque no es problematizadora, ni induce a movilizar conocimientos aplicándolos a una situación real
<b>Organización lógica del contenido</b>	Presenta un contenido conceptual que requiere de la experiencia para dotar de significado al material presentado y poder generalizarlo
<b>Organización metodológica</b>	Requiere del uso de estrategias de clasificación y de procedimientos de categorización y actividades simbólicas y prácticas reales



<b>Organización cognitiva</b>	Implica el uso de procesos simples de observación, reconocimiento, identificación y clasificación
<b>Organización sociocultural</b>	Está ubicado en un contexto extra disciplinar vinculado a situaciones reales, pero que sólo cobran significado para niños de determinados contextos socioculturales

#### **Variables de construcción del reactivo que afectan su resolución**

<b>Objetivo del ítem</b>	El objetivo del ítem de "identificar el significado de una acotación en un plano no está claro, sobre todo porque los símbolos que se usan no están ubicados en planos y además no son de uso común en todo tipo de contexto, más bien parecería que el objetivo del ítem es interpretar símbolos
<b>Expresiones socioculturales del lenguaje</b>	Habría que ver si el vocablo restaurante es accesible y familiar para los niños de los diferentes contextos en los que se aplica la prueba ENLACE, quizá fonda, sería un término más común
<b>Características perceptuales y de redacción de los distractores</b>	En los distractores más que problemas perceptuales, lo que se observa es problemas de contextualización. Los símbolos utilizados no son familiares en los diferentes contextos
<b>Estrategias empleadas para su resolución</b>	Si los alumnos no están familiarizados con estos símbolos, lo más probable es que recurran a la adivinación como estrategia para encontrar la respuesta correcta
<b>Justificación de distractores</b>	Los símbolos utilizados como distractores, pertenecen a distintos conceptos, pero más que un error de conceptualización, lo que pueden reflejar es un error en el manejo del vocabulario

### **3.3. Conceptualización y caracterización de la evaluación de la comprensión lectora en la prueba ENLACE**

El análisis realizado al 30% de los reactivos de comprensión lectora en la Prueba ENLACE, permite acercarnos a la conceptualización de lo que se entiende por comprensión lectora en la prueba ENLACE<sup>68</sup>

El Cuadro 26 presenta la recapitulación de las características de los reactivos que evalúan la comprensión lectora en la prueba ENLACE 2008 de 6° grado

<sup>68</sup> A diferencia de la prueba PISA, donde se define a la comprensión lectora, como el empleo y la reflexión a partir de textos escritos, con el fin de alcanzar las metas propias, desarrollar el conocimiento y el potencial personal y participar en la sociedad" (OCDE, 2001<sup>3</sup>, p.32), la ausencia de una teoría del contenido (Díaz Barriga, 2004) es uno de los problemas fundamentales que de acuerdo con investigaciones previas (Padilla, 2005) se pueden identificar en la evaluación del aprendizaje a través de pruebas a gran escala en México

Cuadro 26

**Características de los reactivos para evaluar la comprensión lectora en la prueba ENLACE 2008 6° grado**

Grado de dif.	No. Reactivo	Tarea	Contenido	Estrategias	Procedimientos	Actividades	Procesos	Contextos
Alta	1	Comprensión	Conceptual	organización	Inferencia	simbólicas	Procesos simples de Ordenamiento Recuperación de Información Relevante Procesos complejos de Síntesis	Vinculado a situaciones reales
	9	Reproductiva	Datos	repaso	repetición	simbólicas	Reconocimiento Identificación	Disciplinar escolar
	98	Comprensión	Conceptual	selección	Paráfraseo inferencia	simbólicas	Proceso complejo de interpretación de textos	Escolar, disciplinar
	147	Reproductiva	Hechos	Repaso Recirculación de la información	acumulación	reproductivas	Simple de reconocimiento e identificación	Extradiisciplinar, vinculado a situaciones reales, pero poco significativas para los estudiantes

Grado de dif.	No. reactivo	Tarea	Contenido	Estrategias	Procedimientos	Actividades	Procesos	Contexto
Medio	39	comprensión	procedimental	Construcción	Elaboración	Prácticas	Proceso simple de ordenamiento, descripción, recuperación de la información	Escolar disciplinar
	41	comprensión	Conceptual	Organización de la información	paráfraseo	Simbólicas	Procesos simples de ordenamiento y recuperación de la información Procesos complejos de síntesis	Escolar extradiisciplinar
	103	comprensión	Conceptual	Selección y clasificación de la información	Identificación de palabras clave categorización	prácticas	Procesos simples de reconocimiento, identificación, ordenamiento, recuperación de la información	Extradiisciplinar Vinculado a situaciones reales
	104	comprensión	Conceptual	Organización de la información	Identificación de palabras clave categorización	Simbólicas prácticas	Procesos simples de reconocimiento, identificación, ordenamiento, recuperación de la información	Extradiisciplinar Vinculado a situaciones reales

Grado de dif.	No. Reactivo	Tarea	Contenido	Estrategias	Procedimientos	Actividades	Procesos	Contexto
Medio	110	comprensión	Conceptual	Selección y organización de la información	Identificación de hechos o ideas	Simbólicas y prácticas	Proceso simple de reconocimiento, identificación, organización y recuperación de la información	Extradisciplinar Vinculado a situaciones reales
	116	comprensión	conceptual	Clasificación de la información	Categorización de hechos	simbólicas	Procesos simples de reconocimiento, identificación y recuperación de la información	Escolar-disciplinar
	142	Reproductiva de ejercitación rutinaria	Hechos	Repaso y recirculación de la información	Repetición	Reproductivas	Procesos simples de observación, reconocimiento e identificación	Escolar-disciplinar

Grado de dif.	No. Reactivo	Tarea	Contenido	Estrategias	Procedimientos	Actividades	Procesos	Contexto
Bajo	10	comprensión	Conceptual	clasificación	categorización	Simbólicas y prácticas reales	Procesos simples de observación, reconocimiento e identificación	Escolar-disciplinar

La concentración del análisis realizado a los reactivos seleccionados, presentada en el Cuadro 26, permite caracterizar la evaluación de la comprensión lectora en la prueba ENLACE como una evaluación centrada en tareas que implican una comprensión, pero que no son problematizadoras, ni promueven la movilización de conocimientos para su aplicación a una situación real. Incluso, se encuentran tareas reproductivas de ejercitación rutinaria y poco significativas.

El contenido a evaluar en la mayoría de los casos, es un contenido meramente conceptual que requiere de establecer interrelaciones en los elementos básicos y que de acuerdo con Taba (1979) si consideramos a los conceptos como sistemas complejos, el aprendizaje de éstos requeriría de la estructuración de experiencias previas para dotar de significado a la información presentada. Aunque también se encuentran contenidos que requieren de un bajo nivel de abstracción, referentes a la adquisición de hechos y datos específicos. La parte procedimental que demanda una acción sistematizada para realizar una tarea o alcanzar un objetivo, es prácticamente nula.

Esto implica la creación de situaciones didácticas o experiencias de aprendizaje que recurran a estrategias de selección, organización y recirculación de la información. Y a la aplicación de procedimientos de repetición, identificación, categorización, realización de inferencias, parafraseo y actividades de tipo simbólicas, prácticas y reproductivas.

Aunque algunos de los reactivos están catalogados con un grado de dificultad alta, los procesos cognitivos implícitos en la realización de las tareas presentadas son procesos simples de reconocimiento, recuperación, identificación y ordenamiento de la información presentada. Sólo en muy pocos casos se recurrió al uso de la síntesis.

Se utilizan contextos tanto escolares de tipo disciplinar, como vinculados a situaciones reales, pero poco significativas para los estudiantes de los diferentes contextos socioculturales.

Lo que da cuenta que la comprensión lectora en la prueba ENLACE, se concibe como una interrelación de ideas abstractas y adquisición de datos y hechos específicos, que requieren de la interrelación de elementos básicos y de la recirculación de la información. Esto requiere de la

creación de situaciones didácticas centradas en la selección, organización y recirculación de la información con procedimientos centrados en la repetición, identificación, categorización de la información y realización de inferencias y parafraseo, que demandan la realización de actividades simbólicas y reproductivas. Lo que implica el uso de procesos simples de reconocimiento, recuperación, identificación y ordenamiento de la información presentada. Para lo cual se recurre a contextos de tipo disciplinar, vinculados a situaciones reales pero con situaciones que no siempre resultan significativas para los niños de los diferentes contextos socioculturales.

Si tomamos en cuenta que de acuerdo con Díaz Barriga, Arceo y Hernández (2002) "La comprensión de textos es una actividad constructiva, compleja de carácter estratégico, que implica la interacción entre las características del lector y del texto, dentro de un contexto determinado" (p. 275). Es decir, para que el lector pueda construir los significados a partir de lo que se sugiere en el texto, requiere tanto del uso de recursos cognitivos pertinentes, como de ciertos esquemas de representación, habilidades y estrategias, como la interpretación y la inferencia. Además la relación con el contexto en que se ubica y la situación social a la que se refiere, desempeña un papel fundamental para la comprensión del texto. De igual manera es fundamental que exista una fuerte vinculación entre las características del texto y las características, cognitivas, y actitudinales del lector, tales como intereses y conocimientos previos.

El Cuadro 27, presenta las características del manejo de la comprensión lectora en la prueba ENLACE y las características que requiere la comprensión de un texto.

Cuadro 27

<b>Conceptualización, componentes y estrategias didácticas de la evaluación de la comprensión lectora en la prueba ENLACE y en el proceso de comprensión lectora</b>		
<b>Evaluación de la comprensión lectora en la Prueba ENLACE</b>		<b>El proceso de comprensión lectora</b> (Díaz Barriga, Arceo y Hernández, 2002, p. 275-277)
Conceptualización	Interrelación de ideas abstractas y adquisición de datos y hechos específicos	Actividad constructiva, compleja de carácter estratégico, que implica la interacción entre las características del lector y del texto, dentro de un contexto determinado

<b>Evaluación de la comprensión lectora en la Prueba ENLACE</b>		<b>El proceso de comprensión lectora</b> (Díaz Barriga Arceo y Hernández, 2002, p. 275 -277)
Componentes	Interrelación de elementos básicos y de recirculación de la información	Recursos cognitivos pertinentes Conocimientos previos o esquemas de representación Habilidades psicolingüística de corte lexical, sintáctico, semántico y pragmático Factores motivacionales Estrategias didácticas
Estrategias didácticas	Selección Organización y recirculación de la información Procedimientos de repetición, identificación y categorización de la información Realización de inferencias y paráfraseo Actividades simbólicas y reproductivas	Reconocimiento de palabras Organización de significados Reconstrucción de relaciones entre proposiciones Extracción de significado Asignación de significado Movilización de conocimientos previos)

La comparación entre las características que presenta la evaluación de la comprensión lectora en la prueba ENLACE y las características que requiere el proceso de comprensión lectora, da cuenta de que la evaluación de la comprensión lectora en la prueba ENLACE, considera básicamente actividades muy elementales de ejecución automática, como reconocimiento, e identificación, a la que autores como Díaz Barriga Arceo y Hernández, (2002) llaman de microprocesamiento. Sin embargo, para estos autores la complejidad que implica el proceso de comprensión lectora, reside en la profundidad de los diversos niveles en que ésta se realiza y requiere de actividades tanto de microprocesamiento como de macroprocesamiento. Estas actividades de macroprocesamiento, son las que se realizan de manera más consciente y tienen que ver con la representación semántica de la naturaleza abstracta y global del texto. Requieren de la jerarquización, de la aplicación de inferencias basadas en conocimientos previos, así como de la construcción de un modelo situacional, es decir de un "mundo creado para el texto" (Díaz Barriga Arceo y Hernández, 2002, p. 278), lo cual representa el nivel más complejo y determinante para la comprensión de un texto.

De acuerdo al análisis realizado, vemos que la estructura de los reactivos de la prueba ENLACE para evaluar la comprensión lectora, difícilmente promueven el macroprocesamiento de la

información, sobre todo porque no consideran la movilización de conocimientos previos y el empleo de contextos cercanos a la realidad de todos los estudiantes, que permitan la construcción de modelos situacionales. Lo que hace pensar que aunque se utilizan reactivos con diferente grado de complejidad, la evaluación de la comprensión lectora en la prueba ENLACE, se queda todavía en un nivel muy básico.

Hasta aquí, tenemos que la información que más se da a conocer de la prueba ENLACE, es de tipo descriptivo, técnico, operativo y normativo, como si la evaluación del aprendizaje fuera solamente un asunto de orden mecánico.

A pesar de la publicación del texto "Sugerencias para el uso pedagógico de los resultados de ENLACE" (SEP, 2008c), la información que aquí se presenta, consiste en una mera descripción del contenido y propósito del reactivo. Hasta el momento no se cuenta con información referente a los elementos de orden cognitivo, metodológico y sociocultural, implícitos en todo proceso de enseñanza-aprendizaje y evaluación.

El análisis realizado permite dar cuenta de algunos de los alcances, limitaciones y sentido que esta prueba adquiere como instrumento de evaluación de la educación primaria. En especial, como instrumento de evaluación de la comprensión lectora, que es el área que se tomó como referencia.

### **3.4 Alcances, limitaciones y sentido de la estructura didáctica de los reactivos para evaluar la comprensión lectora en la prueba ENLACE 2008**

El análisis realizado, permite dar cuenta de ciertos alcances y limitaciones que desde una mirada didáctica pedagógica se perciben en esta prueba.

La evaluación de la comprensión lectora en la prueba ENLACE, se hace a partir de multireactivos, lo que proporciona la evaluación integrada de diversos conocimientos.

Se utilizan diferentes tipos de textos, para evaluar la comprensión lectora, lo que permite diferenciar las funciones de la lectura. Sin embargo los tipos de textos que se utilizan no siempre resultan

cercanos y familiares a los diversos contextos en los que se encuentran los estudiantes, lo que dificulta su comprensión

Aunque se trata de incluir reactivos con diferente nivel de dificultad, en los reactivos analizados, predomina el uso de procesos cognitivos simples, que poco o nada tienen que ver con los requeridos para una comprensión lectora. No se incluyen procesos cognitivos complejos. Lo que lleva a pensar que aunque se le da un tratamiento conceptual a la evaluación de la comprensión lectora, ésta sigue concibiéndose como algo mecánico y rutinario.

En cuanto a su función didáctica, el contenido de los reactivos analizados no representan una oportunidad para demostrar lo aprendido, porque no movilizan conocimientos previos. Además no representan un medio para corregir el error, porque los distractores utilizados no ayudan a identificar el error en el manejo del contenido. Ante lo cual, se puede decir que este examen continúa siendo un instrumento de verificación, más que un medio para ayudar a aprender y demostrar lo aprendido, que desde una perspectiva didáctica, sería la función que éstos tendrían.

En relación a las variables de construcción del reactivo que afectan su resolución, se observaron algunos problemas como los siguientes:

- Falta de claridad y definición de los objetivos del ÍTEM, por la redacción de algunos de los reactivos.
- El tipo de vocabulario empleado, favorece de manera desigual el éxito de los niños, porque los códigos sociolingüísticos, no adquieren el mismo significado, para los niños ubicados en contextos diferentes a los trabajados.
- En cuanto a la redacción de los distractores, no se encontraron muchos problemas perceptuales, aunque sí de redacción.
- Las características que presentan los reactivos en cuanto a su construcción, inducen a que los alumnos acudan a la eliminación o adivinación como estrategias para dar con la respuesta correcta.



- En cuanto a los distractores empleados, se vio que éstos no reportan un determinado error en el manejo del contenido, lo que cierra la posibilidad de promover un aprendizaje a partir del error, que sería una de las funciones didácticas de las pruebas

Con el fin de tener una visión general de la prueba, en el Cuadro 28, se presenta una síntesis de los alcances y limitaciones identificadas en los diferentes aspectos que abordan la prueba ENLACE

Cuadro 28

**Alcances y limitaciones de la prueba ENLACE**

<b>Aspectos que comprende</b>	<b>Alcances</b>	<b>Limitaciones</b>
Información con la que se cuenta	Mayor información en su proceso de construcción y tipificación	Pero no sobre su proceso didáctico
Normas para su elaboración	Se establece como norma de elaboración, la definición del dominio de aplicación	No cuenta con una definición del objeto a evaluar
Contextualización de la tarea	Empleo de textos diversos y uso de contextos reales en multirreactivos	Situaciones y textos no siempre cercanos a la realidad de los niños
Nivel de dificultad	Reactivos con diferente nivel de dificultad	Uso de procesos cognitivos simples, no suficientes para una comprensión lectora. No se incluyen procesos complejos
Evaluación de la comprensión lectora	Tratamiento conceptual de la comprensión lectora	Se concibe a la comprensión lectora como algo mecánico y rutinario

Aspectos que comprende	Alcances	Limitaciones
Variables que afectan su resolución		Falta de claridad y definición de los objetivos del ÍTEM Los códigos sociolingüísticos, no adquieren el mismo significado, para los niños ubicados en contextos diferentes Las características que presentan los reactivos en cuanto a su construcción, inducen a que los alumnos acudan a la eliminación o adivinación como estrategias de solución Los distractores utilizados, no ayudan a identificar el error en el manejo del contenido
Función		Mero instrumento de verificación No representa una oportunidad para demostrar lo aprendido
Sentido		Origen político Tiene como base la rendición de cuentas Centrada en aspectos técnicos No presenta un abordaje didáctico

Con base en lo que se presenta en el Cuadro 28, podemos decir que la prueba ENLACE, como instrumento de evaluación, es una prueba cuyo contenido, dista mucho de los pretendido y establecido en los planes y programas de estudio en cuanto a habilidades y tipos de aprendizaje a desarrollar. Uno de sus propósitos centrales del Plan 1993, es el desarrollo de habilidades entre éstas la comprensión lectora, de acuerdo al análisis realizado no hay tal desarrollo de habilidades. Las tareas presentadas, requieren el uso de procesos básicos, no suficientes para desarrollar la comprensión lectora. En los reactivos analizados, encontramos un contenido más de tipo mecánico y abstracto, que significativo; éste requiere de estrategias y procedimientos didácticos memorísticos y rutinarios.

Lo que lleva a pensar por una parte si éste es el tipo de aprendizaje y enseñanza con la que se pretende elevar la calidad de la educación básica? Y por otra, qué tan válido es enjuiciar el trabajo de los docentes, a través de instrumentos de evaluación que presentan ciertas limitaciones en cuanto a su contenido y requerimientos didácticos y que además no corresponden a las formas de trabajo del aula y tipo de aprendizajes que ahí se requieren.

## CONCLUSIONES

---

La investigación que presentamos, inscrita en el debate de la evaluación externa como sinónimo de calidad educativa, revela importantes problemas de índole político-social y educativo, que pruebas a gran escala como ENLACE, han generado en nuestro sistema de educación y en especial en nuestros docentes y estudiantes

La implementación de la prueba ENLACE en nuestro país y la importancia que ésta ha cobrado para juzgar la efectividad de todo un sistema y en especial del trabajo de docentes y alumnos, responde a intereses político-administrativos, más que educativos. De hecho su creación y ejecución obedeció más a una coyuntura gubernamental, que a una preocupación por ayudar a mejorar lo que sucede en nuestras escuelas. La prueba ENLACE ha funcionado sólo como un medio para ejercer un control sobre docentes y escuelas. Hasta el momento no se tienen indicios de que la tan pregonada mejora de la calidad educativa, que justifica su creación, se haya podido vislumbrar.

A pesar de que el examen en sus orígenes dentro de la didáctica, formaba parte del método de enseñanza, la estandarización que de éstos se ha hecho y el carácter cientificista que han ido adquiriendo al sustentarse en una teoría del test, ha propiciado un manejo de tipo gestivo, donde lo que se privilegia es la función social de tipo certificativa de la evaluación. Los exámenes a gran escala hoy en día, están impregnados por cuestiones de hegemonía, diferenciación y legitimación, más que de mejora educativa. Esto deja ver la necesidad de hacer que el examen, más que un mero instrumento para medir el rendimiento escolar, sea aprovechado por el docente como un medio para revisar y apoyar la actividad del aula y promover mejores aprendizajes en sus alumnos.

El eje central de este trabajo, se centra en proponer un tratamiento más pedagógico para el estudio de las evaluaciones estandarizadas, con el propósito de recobrar el sentido didáctico del examen como instrumento de evaluación.

Un balance de los alcances y limitaciones de los sistemas de evaluación estandarizada y la importancia que las acciones de evaluación externa han cobrado en la actualidad, nos lleva a

plantear que existen algunos logros que se pueden destacar, pero también serían dificultades, relacionadas sobre todo con el contenido y manejo que a éstos se les ha dado

Reconocemos como avances de estos sistemas de exámenes, que mientras hace algunos años, el debate educativo estaba centrado principalmente en la cobertura y el acceso a la educación básica, actualmente, la preocupación está abocada a conocer, o más bien *medir* los logros manifestados por los estudiantes. También podemos decir que hay un interés en toda la región latinoamericana por una mayor transparencia en cuanto a la difusión de resultados, así como por mejorar la capacidad técnico-metodológica para la construcción de pruebas y por hacer más preciso el procesamiento de los datos. De igual manera es clara la importancia que se le da, a indagar los factores que inciden sobre los resultados obtenidos.

Existen defensores de estas pruebas, que sostienen la existencia de ciertas bondades en materia de rendición de cuentas. Lo que da pie a la adopción de la evaluación externa con mayor legitimación.

Sin embargo podemos identificar en los avances enunciados, problemas importantes que requieren ser atendidos.

Una de las primeras dificultades que encontramos en especial en la prueba ENLACE, es la del contenido. Porque el logro que a través de éstos se pretende medir, se refiere sólo a poder resolver adecuadamente ciertos reactivos o inclusive a "atinarle" a la respuesta establecida como correcta, de acuerdo con criterios establecidos y respuestas prefabricadas. Este tipo de exámenes no evalúan lo que el estudiante realmente sabe, ni las habilidades que ha conseguido desarrollar, ni lo que puede o no puede hacer. Como diría Gil (2010), "los exámenes no miden el mérito, sólo miden la ambición y la disposición a obedecer, también, la habilidad para someterse al examen".

Por otra parte, la difusión y uso de resultados que de éstos se hace, presenta serias dificultades también, porque no basta con informar, sino que el problema está en lo que se informa, en cómo se informa y el para qué se informa.

Así, desde una mirada pedagógica, se perciben problemas importantes en el contenido de este tipo de exámenes, en la manera en como se difunden los resultados y en lo que éstos generan

Lo anterior ha permitido destacar requerimientos fundamentales a los que este trabajo ha intentado dar respuesta

Nuestro planteamiento parte de que una de las situaciones más sobresalientes, sin duda reside en la necesidad de un abordaje didáctico de los exámenes a gran escala, que incluya un análisis del contenido mismo del examen, de su vinculación con ciertas estrategias metodológicas y del tipo de aprendizaje que promueve

El empleo que hasta el momento se les ha dado a los exámenes a gran escala, revela que éstos han quedado aislados del método didáctico, porque no acompañan el proceso de enseñanza-aprendizaje, ni mucho menos permiten observar dificultades en el aprendizaje para hacer una construcción y reconstrucción del proceso de tanto de enseñanza como de aprendizaje del estudiante. La información que de ellos se obtiene sólo da a conocer si el alumno aprobó y llegó al estándar establecido, o si ha adquirido, a través de los manuales que para ello se han creado, la suficiente habilidad para "aprender" a contestar exámenes, mas no permite conocer el tipo y nivel de aprendizaje que el estudiante ha alcanzado y la manera de ayudar a que *todos lleguen al aprendizaje*. El principio didáctico de "enseñar todo a todos" del viejo Comenio (1922/2000), ha quedado fuera de la escuela y es totalmente ajeno a la aplicación de este tipo de pruebas, cuya función principal es certificar y discriminar al que puede, del que no puede

En el plano de lo didáctico y con un enfoque de evaluación formativa, un examen tendría la función de retroalimentar el trabajo del docente en el aula. Sin embargo lo que de éstos se obtiene son sólo reportes técnicos, que en nada ayudan al profesor a saber qué es lo que él y sus alumnos tienen que mejorar. Al docente se le juzga, valora y clasifica por los resultados obtenidos por sus alumnos, pero no se le ayuda a ver la pertinencia y el por qué de esos resultados

Y aunque sabemos que la evaluación por sí sola no mejora el aprendizaje y no es a través de pruebas a gran escala que se va a lograr la tan ansiada calidad educativa. Más que ayudar a un

manejo más óptimo del proceso enseñanza-aprendizaje, la ausencia de un tratamiento didáctico en estos sistemas de exámenes, ha empobrecido el trabajo del aula y el aprendizaje de los estudiantes. Por el tipo de contenido que en éstos se maneja, por la visión de conocimiento y aprendizaje que transmiten y por la alteración que causan en los procesos escolares, al centrar como en antaño, la acción de docentes y alumnos en torno a la prueba. Tal parece que ahora lo importante no es aprender y desarrollar capacidades, sino pasar la prueba. Se le da más importancia a los "manuales para contestar exámenes" que a los libros de texto o al programa escolar. Las pruebas han venido a constituir más que una ayuda para el docente, un obstáculo y una amenaza. Ante la premura de la "prueba" y de la "semana de evaluación", ya no hay tiempo para pensar en crear y recrear la enseñanza y en generar las estrategias necesarias que despierten en sus alumnos el gusto por aprender. Como diría Ramírez (2010), el sentido de la educación se ha perdido con este sistema de exámenes.

Esta búsqueda, por recobrar el sentido didáctico de los exámenes, sobre todo en su modalidad a gran escala y hacer del examen un instrumento más educativo que judicial, nos lleva a hacer una propuesta de análisis que permite un empleo más didáctico de los exámenes a gran escala. Nuestra intención es devolverle el sentido y significado educativo a un examen y que a través del análisis de los reactivos de los exámenes a gran escala que actualmente se están aplicando a nuestros estudiantes, se pueda transitar de una mera información, a una orientación y devolución de resultados con un sentido formativo.

A través del modelo propuesto en este trabajo, donde se presenta un análisis tanto conceptual, como metodológico y sociocultural de los reactivos de una prueba, buscamos hacer del examen más que un regulador de la enseñanza, una herramienta más del docente, que le ayude a conocer el tipo de aprendizaje que sus alumnos han logrado desarrollar y con base en esto, promover formas más significativas y apasionadas de aprender.

Sin embargo también queremos destacar, que la construcción del modelo que aquí se propone, para el análisis de los reactivos de un examen a gran escala con una perspectiva didáctica, no resultó una tarea fácil. Su construcción y aplicación se volvió algo complejo. Sobre todo porque nos encontramos con la dificultad que existe para dar un tratamiento didáctico a modelos de evaluación

construidos desde una lógica ciferente. Es decir, frente a pruebas estandarizadas que tratan de homogenizar y discriminar, tenemos que la didáctica parte del principio de dar respuesta a las necesidades individuales para facilitar el aprendizaje. La teoría de la medición está sustentada en el principio de diferenciar al que sabe, del que no sabe, y el propósito de la didáctica, es hacer que todos aprendan.

Tenemos entonces que la intencionalidad de las pruebas a gran escala y de la didáctica son opuestas. La didáctica en un sentido amplio, pretende mejorar las condiciones de desarrollo de los educandos, mientras que las pruebas a gran escala constituyen instrumentos de diferenciación social, que lo que buscan es la clasificación y el control de los sujetos. La complejidad de las pruebas estandarizadas y sus sustentos teóricos-conceptuales, no responden a una perspectiva didáctica, porque obedecen a tradiciones y finalidades distintas.

De ahí que ante el desencuentro existente entre los planteamientos didácticos y los sustentos conceptuales de las pruebas a gran escala. Así como a la escasez de estudios didácticos en el ámbito mexicano,<sup>89</sup> creemos importante buscar una alternativa para construir un puente entre estas dos tradiciones, que aunque obedecen a una racionalidad distinta, dada la importancia que las pruebas a gran escala han adquirido en los sistemas educativos, la posibilidad de establecer un diálogo entre lo didáctico, lo psicológico y lo psicométrico, se ha vuelto algo obligado para los estudiosos de la educación y de la evaluación del aprendizaje. Ante lo cual consideremos que un punto de encuentro entre estas dos lógicas, sería el aprendizaje y los procesos mediante los que éste se desarrolla, al ser dos disciplinas que se desarrollan entre sí.

Es precisamente por la inexistencia de este vínculo entre la didáctica y la evaluación estandarizada, que motivados por las debilidades encontradas en el uso que se les está dando a los exámenes a gran escala y el impacto social que éstos han tenido, buscamos a través de la construcción y aplicación del modelo de análisis de reactivos que proponemos, ver con una mirada más didáctica a los exámenes a gran escala.

---

<sup>89</sup> Cabe destacar la desaparición de la didáctica en los planes de estudio de las normales, suscitada a partir de la reforma curricular de la Licenciatura en Educación Primaria ocurrida en 1997.

El modelo propuesto, presentó dificultades para su construcción y aplicación, consideramos que ayuda a darle un sentido diferente a los exámenes estandarizados, al vincularlos más que con resultados cuantitativos, con procesos de aprendizaje y de enseñanza

Decidimos aplicar el modelo de análisis de reactivos que elaboramos, a una de las pruebas que mayor impacto social ha causado en nuestro sistema educativo, como es la prueba ENLACE Prueba que desde el 2006, se ha venido aplicando de manera consecutiva año tras año en nuestro país durante el mes de abril Sin embargo cabe preguntarnos, si después de estos cinco años, en que los estudiantes de educación básica, han sido sometidos a exámenes de este tipo, nuestro sistema educativo y más aún nuestro sistema de enseñanza, ha mejorado o se ha empobrecido

Lo que si podemos decir es que, la aplicación del modelo propuesto, permite dar cuenta de varios problemas de los exámenes a gran escala como acción para evaluar el aprendizaje y de la prueba ENLACE en particular

Si consideramos que en el debate didáctico, un principio fundamental, es la estrecha relación que debe existir entre contenido y método, así como entre formas de aprendizaje, estrategias de enseñanza y de evaluación El análisis realizado de algunos de los reactivos de español de la prueba ENLACE permite ver c la manera en que un examen nacional como la prueba ENLACE, concibe a la comprensión lectora y a partir de una caracterización de la evaluación que de ésta se hace, se puede encontrar la contraposición que existe entre las características que requiere el desarrollo de una comprensión lectora y la manera en como la prueba ENLACE la evalúa De acuerdo a los resultados obtenidos en el análisis realizado, tenemos que en lugar de favorecer el aprendizaje y desarrollo de la comprensión lectora, lo limitan, porque los contenidos, procesos y contextos involucrados en su evaluación, son de un nivel muy elemental que difícilmente promueven la representación semántica y la construcción de modelos situacionales de representación que ayuden a una mejor comprensión lectora.

El análisis realizado a algunos de los reactivos de comprensión lectora de la prueba ENLACE de 6° grado, deja ver la limitación existente en cuanto a contenido, procesos cognitivos y estrategias metodológicas a las que esta prueba recurre para la evaluación de la comprensión lectora



Tenemos entonces que más que promover un mejor aprendizaje, potencializando el desarrollo de los procesos que se requieren, la valoración y por lo tanto el trabajo de la comprensión lectora, de acuerdo a lo solicitado por la prueba ENLACE, se restringe a cuestiones mecánicas y sin sentido

Por otra parte, el análisis de las variables implícitas en la construcción de los reactivos, revelo serios problemas en cuanto a elaboración misma de la prueba

Lo que hace ver a la prueba ENLACE, como un examen a gran escala, que además de presentar dificultades en su construcción, difícilmente puede ayudar a promover una mejora de la enseñanza y el aprendizaje de la comprensión lectora en la educación primaria

Aunque la intención de exámenes como ENLACE, no es de tipo didáctico, se dice que uno de los propósitos principales a las que obedece su aplicación, es la mejora de la educación. Cabe preguntar entonces, si una prueba centrada en contenidos mecánicos de corte abstracto, que requiere el uso de procesos memorísticos y rutinarios y que además dista mucho de lo que se plantea en los planes y programas de estudio en los que se basa, puede ayudar a promover procesos de enseñanza y aprendizaje que ayuden a mejorar el desarrollo cognitivo y social de nuestros estudiantes

Todo esto, lleva a pensar en el riesgo que representa, enjuiciar el trabajo de docentes y determinar el aprendizaje de los alumnos, a través de pruebas a gran escala como la prueba ENLACE, que presenta serias dificultades tanto en su construcción, como en lo que evalúa y reporta. Lo que a su vez limita la acción didáctica del docente en lugar de promoverla

Es importante advertir además sobre el empobrecimiento que en materia de aprendizaje están generando exámenes como ENLACE. El carácter judicial que han mantenido y la influencia que han venido ejerciendo para valorar la situación de la educación en nuestro país y juzgar el trabajo del docente, ha ocasionado que las actividades del aula se centralicen en revisar y trabajar los contenidos de ENLACE, más que en promover aprendizajes significativos

Tenemos entonces, que la importancia que pruebas a gran escala como ENLACE han cobrado en nuestro país por el impacto social, político y administrativo que estas han generado, determina en gran parte no sólo lo que se enseña y aprende, sino como se enseña y aprende en nuestras escuelas. Ahora lo más importante en nuestro sistema educativo, ya no enseñar y recrear el aprendizaje, sino preparar y adecuar lo que se enseña a lo que se pide en una prueba.

Pruebas a gran escala como ENLACE, han provocado un cambio en el sentido de la enseñanza, el aprendizaje y la evaluación del aprendizaje. El examen, más que un medio para revisar la enseñanza, se ha cristalizado hoy, como el único fin de la enseñanza. El aprendizaje más que una cuestión de desarrollo individual y social, se ha convertido en un asunto de homogenización y estandarización. La evaluación más que una manera de demostrar lo aprendido, es hoy en día una manera de limitar y controlar lo que se enseña y aprende.

Éstos son los mensajes implícitos que la evaluación a gran escala a través de pruebas como ENLACE, están generando en materia educativa, tanto a docentes, como a alumnos y sociedad en general.

#### **Consideraciones finales**

Si quisiéramos puntualizar en cuáles son los mayores aportes del estudio que presentamos, tendríamos que éstos se pueden enunciar en dos sentidos. En un plano amplio, la investigación realizada, deja ver la necesidad de un mejor desarrollo de sistemas de evaluación a gran escala, o la modificación de los existentes. Lo que requiere de una seria revisión sobre ¿qué es lo que se puede o debe evaluar a través de este tipo de pruebas? ¿cuál es el sentido de estas evaluaciones?, ¿cuáles son sus usos, propósitos y funciones? y ¿qué tipo de desafíos conceptuales, metodológicos y sociológicos plantean? Así como la búsqueda de los mecanismos y criterios más apropiados para el uso y difusión de los resultados, poniendo un énfasis especial en lo que se informa y a quién se informa. De tal manera que efectivamente sirvan para mejorar la enseñanza y el aprendizaje en nuestras aulas de educación básica. Y de manera particular, el modelo propuesto responde a la necesidad de que a partir de un análisis didáctico de los test, se pueda impulsar el trabajo escolar.

Sin embargo, una limitación importante que le vemos al estudio realizado y que se presenta como opción para futuras investigaciones, es conocer y analizar la voz de los actores y de los principales destinatarios de estos exámenes, (docentes y alumnos) con respecto al impacto que estos exámenes generan tanto en la enseñanza como en el aprendizaje. Porque una preocupación importante que se deriva de los análisis realizados es ¿Qué piensan los profesores de educación básica con respecto a lo que estos exámenes están generando tanto en su función como docentes, como en el sistema educativo en general? ¿Qué genera en los estudiantes, la importancia que este tipo de exámenes ha cobrado? ¿Cuál es su impresión de lo que está sucediendo? ¿En qué manera les afecta o favorece? ¿Cuáles son sus estrategias de sobrevivencia?

Muchas son las interrogantes que quedan por resolver y muchos más los problemas de índole pedagógico a los que la evaluación de tipo estandarizada tiene que responder. Si bien se reconocen ciertos adelantos de tipo técnico, en materia de construcción de pruebas, procesamiento y difusión de los resultados de un examen a gran escala, la situación con respecto al tipo de contenido que este tipo de pruebas evalúa y al aprendizaje que promueven, así como de la interpretación pedagógica de los resultados, todavía deja mucho que desear. La finalidad de los test, sigue siendo demostrar el fracaso escolar y no generar las acciones necesarias para impulsarlo y promoverlo.

Queda como reflexión final la evaluación del aprendizaje escolar a través de pruebas a gran escala, ¿debe seguir promoviendo la diferenciación social y educativa?, o podemos hacer algo por encontrarle un sentido diferente.

## REFERENCIAS

- Aboites, H (2000) "Examen único y su cultura de la evaluación en México" En A Díaz Barriga y T. Pacheco (coord) *Evaluación Académica México* Centro de Estudios sobre la Universidad y la Educación (CESU) / Fondo de Cultura Económica (FCE)
- Aboites, H (2006) "El CENEVAL, monopolio privado y comercial de evaluación" *Vanguardia Proletaria* Recuperado 8 de abril 2006, de <http://cbbcxixi.spaces.live.com/blog/cns7118F842E33981361194.entry>
- Aebli, H (2001) "Actividades vitales y actividades de aprendizaje" En *Factores de la enseñanza que favorecen el aprendizaje autónomo*. Madrid Narcea
- Aebli H (2002) *Doce formas básicas de enseñar. Una didáctica basada en la psicología de Jean Piaget*. Madrid Narcea
- Allabach, P (2005) "Los dilemas de los rankings" (Tr Salvador Medina) *Campus Milenio*. Suplemento universitario No 155 Recuperado el 28 de noviembre de 2005, de <http://www.campusmilenio.com>
- Airasian, J (2001) *Classroom assessment. Concepts and applications* Boston McGraw Hill
- Anastasi, A y Urbina S (1998) "Pruebas estandarizadas de aprovechamiento" En *Test Psicológicos*. (7ª Ed) México Prentice Hall
- Administración Nacional de Educación Pública (ANEP) (2002a) *Los niveles de desempeño al inicio de la educación primaria. Estudio de las competencias Lingüísticas y Matemáticas. Primer informe*. Uruguay, Montevideo Administración Nacional de Educación Pública (ANEP) Consejo Directivo Central (CODICEN) Gerencia General de Planeamiento y Gestión Educativa Gerencia de Investigación y Evaluación
- ANEP. (2002b) *Los niveles de desempeño al inicio de la educación primaria. Estudio de las competencias lingüísticas y matemáticas. Evaluación Nacional en el Primer nivel de la Escolandad inicial 5, 1er y 2do año. Primer informe*. Montevideo, República Oriental del Uruguay Rosgal
- ANEP (2004) *La evaluación de la "Cultura científica" en PISA 2003. Marco conceptual y actividades de prueba*. Montevideo, Uruguay ANEP / Gerencia de Investigación y Evaluación.
- ANEP. (2006) *Resolución de problemas para el mundo de mañana. PISA 2003. Medición de competencias transversales*. Montevideo, Uruguay ANEP / Departamento de Investigación y Estadística (DIE)

- ANEP (2007a) *Programa PISA Uruguay* Montevideo, Uruguay CODICEN-DIE
- ANEP (2007b) *Uruguay en PISA 2006. Primeros resultados de Ciencias, Matemáticas y Lectura del Programa Internacional de Evaluación de Estudiantes*. Uruguay, Montevideo CODICEN-DIE
- ANEP (2007c) *Evaluación nacional de aprendizajes en lenguaje y matemática. 6° año de enseñanza primaria -2005*. Uruguay, Montevideo CODICEN-DIE
- ANEP (2008) *Evaluación piloto autónoma. Primer año de educación primaria. Lectura y escritura*. Uruguay, Montevideo CODICEN-DIE
- Ayer, A. J. (1965) *El positivismo lógico*. México Fondo de Cultura Económica
- Backhoff, E y Tirado, F (1992) "Desarrollo del Examen de Habilidades y Conocimientos Básicos" *Revista de la Educación Superior* Julio-Septiembre Vol 6 Num 83 México ANUIES
- Backhoff, E (1998) "Evaluación asistida por computadora, usos recientes y tendencias futuras" En Castañeda S (coord) *Evaluación y fomento del desarrollo intelectual en la enseñanza de las ciencias, artes y técnicas. Perspectiva internacional en el umbral del siglo XXI*. México Porrúa, CONACYT, Facultad de Psicología-UNAM
- Backhoff, E y Rosas, M (2002) "Sistema computarizado de exámenes" En A Bazán A y A Ferrer, A (ed) *Estrategias de evaluación y medición del comportamiento en Psicología*. México Instituto Tecnológico de Sonora (ITS) Universidad Autónoma de Yucatán (UADY)
- Backhoff, E (2003) *Desarrollo, validación e implementación de un sistema para la selección de estudiantes en la Universidad Autónoma de Baja California* Tesis de doctorado México Universidad Autónoma de Baja California Versión en CD [2009, 4 de abril]
- Bain, K (2006) *Lo que hacen los mejores profesores de* Universidad Valencia, España Palacios
- Baudillo, L. (2008) *La evaluación decretada*. Bogotá, Colombia Ediciones SEM
- Benedito, G (1986) "El problema de la medida en psicología" y "El método de los test e inventarios" En Braustein, N., Pasternac, M., Benedito, G, y Sall, F. *Psicología, ideología y ciencia*. (12° Ed) México Siglo XXI
- Bernstein, B (1985) "Clases sociales, lenguaje y socialización". En. *Revista Colombiana de Educación, num. 15*, Bogotá, Colombia, pp 25-44
- Bertoni, A Poggi, M y Teobaldo, M (1996) *Evaluación, nuevos significados para una práctica compleja*. Buenos Aires, Argentina Kapeltz
- Bloom, B (1971) *Taxonomía de los objetivos de la educación*. Argentina El Ateneo

- Bourne, L. Dominowski, R. Loftus, E. y Healy, A. (1986) *Cognitive Proceses. Second edición* Unite States América Prentice Hall
- Bustamante, G y Díaz L. (2003) *Factores asociables al desempeño de los estudiantes*. Bogotá Universidad Nacional de Colombia
- Bustamante, G y Caicedo (2005) *La evaluación: ¿Objetiva o construida?* Colombia, sociedad Colombiana de pedagogía
- Casas, M (2003) "The Use of Standardized Tests in Assessing Authentic Learning—A Contradiction Indeed" *Teachers College Record*, Date Published October 12 de 2006, de TTU <http://www.tcrecord.org> ID Number 11211, Date Accessed 1/20/2006 9 24 48 PM
- Castañeda, S (1998) *Evaluación y fomento del desarrollo intelectual en la enseñanza de las ciencias, artes y técnicas. Perspectiva internacional en el umbral del siglo XXI*. México Porrúa, CONACYT Facultad de Psicología
- Castañeda, S Lugo, E Pineda, L y Romero, N (1998) "Estado del arte de la evaluación y el fomento del desarrollo intelectual en la enseñanza de las ciencias, artes y técnicas" En S Castañeda *Evaluación y fomento del desarrollo intelectual en la enseñanza de las ciencias, artes y técnicas, Perspectiva internacional en el umbral del siglo XXI*. México Porrúa, CONACYT-Facultad de Psicología
- Castañeda, S (2002a) "Explorando la utilidad de los análisis clásicos y de Rasch en la determinación de la consistencia interna de subpruebas cortas de homogeneidad" En A. Bazán y A. Ferrer *Estrategias de evaluación y medición del comportamiento en psicología*. México ITS UADY
- Castañeda, S (2002b) "Diagnostividad e interpretabilidad Retos a la medición de resultados de aprendizaje" En. A. Bazán A y A. Ferrer, A. (ed) *Estrategias de evaluación y medición del comportamiento en Psicología*. México Instituto Tecnológico de Sonora (ITS) Universidad Autónoma de Yucatán (UADY)
- Castañeda, S (2005) *Evaluación del aprendizaje: Investigación y aplicaciones prácticas para el sistema universitario*. Versión en CD
- Castañeda, S (coord) (2006) *Evaluación del aprendizaje en el nivel universitario*. México UNAM – CONACYT.
- Centro Nacional para la Evaluación de la Educación Superior (CENEVAL), (2006) *Informe de actividades 2005*. México CENEVAL
- Comenio, J. A (1922/2000) *Didáctica Magna*. (10° Ed) México Porrúa
- Contreras, L. A. y Backhoff, E (2004) "Metodología para elaborar exámenes criteriosales alineados al currículo" *Documento interno de trabajo*. México INEE

- Contreras, L. A., Backhoff, E. y Larrazolo, N. (2004) "Análisis reticular. Manual para el comité diseñador del examen" *Documento interno de trabajo*. México INEE
- Chadwick, C. (s/f) "La tecnología educativa en América Latina en la década de los 70" Recuperado 15 de enero 2007, de <http://ddd.uab.cat/pub/educar/0211819Xn3p129.pdf>
- Childs, R. (1990) "Gender bias and fairness" *Practical Assessment, Research & Evaluation*, 2(3) Recuperado 8 de mayo de 2008, de <http://PAREonline.net/getvn.asp?v=2&n=3>
- Cimbricz, S. (2002,) "State-mandated testing and teachers' beliefs and practice". *Education Policy Analysis Archives*, 10(2) Recuperado enero 9 de 2007, de <http://epaa.asu.edu/epaa/v10n2.html>
- Coll, C., Pozo, J. I., Sarabia, B. y Valls, E. (1992) *Los contenidos en la reforma. Enseñanza y aprendizaje de conceptos, procedimientos y actitudes*. Madrid: Santillana
- Coll, C. (2003) "Las prácticas de evaluación Una oportunidad para enseñar y aprender" Ponencia presentada en el VII Congreso Galego Portugués de Psicopedagogía Universidad de Coruña y Universidad de Minho Recuperado 7 de mayo de 2008, de <http://www.psvd.edu.es/grintie>.
- Coll, C. (2006) "Cada vez que ha habido cambio curricular, ha sido como resultado de un debate ideológico y no el análisis de las evaluaciones" Entrevista a Cesar Coll. *Docencia*. No. 29
- Contreras Niño, L. A. y Backhoff, E. (2004) "Metodología para elaborar exámenes criterios alineados al currículo" *Documento interno de trabajo*. México INEE
- Contreras Niño, L. A., Backhoff, E. y Larrazolo, N. (2004) "Análisis reticular Manual para el Comité Diseñador del Examen" *Documento interno de trabajo*. México INEE
- Cremin, L. A. (1969) *La transformación de la escuela* Argentina Editorial Bibliográfica Argentina
- Dalmazio, M. y Passos, L. (2001) "Avaliação escolar Desafios e Perspectivas" Em *Ensinar a ensinar*. Didática para a Escola Fundamental e Média Brasil. Thompson
- Darling, L. (2003) "Standards and Assessments Where We Are and What We Need" *Teachers College Record*, February 16, 2003 <http://www.tcrecord.org> ID Number 11109, Date Accessed 12/29/2005 5:12:26 PM
- Day, C. (2006). *Pasión por enseñar. La identidad personal y profesional del docente y sus valores*. Madrid Narcea
- De Alba, A., Díaz Barriga, A. y Viesca, M. (1984). "Evaluación Análisis de una noción" En *Revista Mexicana de Sociología*. Año XLVI. No. 1, enero-marzo
- Del Valle (2008, 17 de mayo) "Preven que el bono mejore a maestros" *Reforma*. Recuperado 17 mayo 2008, de <http://www.reforma.com/nacional/articulo/442882647/>

- Denyer, M. Furnémont, J. Poulain, R. y Vanloubbeeck G. (2007) *Las competencias en la educación. Un balance*. México Fondo de Cultura Económica
- Dewey, J. (1922) "Education as politics" En *Middle works of John Dewey*. Carbondale, Southern Illinois University Press, 1983, Vol. 13.
- Dewey, J. (1958) *Experiencia y educación*. Buenos Aires, Argentina. Losada
- Dewey, J. (1993) *Cómo pensamos. Nueva exposición de la reflexión entre pensamiento reflexivo y proceso educativo*. Barcelona Paidós.
- Dewey, J. (1995-2004) *Democracia y Educación* (6° Ed) Primera edición 1995 Publicado originalmente en 1916. Madrid Morata
- Díaz Barriga, A. (1982) "Tesis para una teoría de la evaluación" En *Perfiles Educativos* No 15, México CESU/UNAM, pp 16-37
- Díaz Barriga, A. (1987) "Problemas y retos en el campo de la evaluación educativa" En *Perfiles educativos* No 37 México CESU/UNAM, pp 3-15
- Díaz Barriga, A. (1993a) *El examen, textos para su historia y debate*. México, CESU / Plaza y Valdés / UNAM (Primera reimpresión, 2000)
- Díaz Barriga, A. (1993b) "El examen un problema de historia y sociedad" En *El examen, textos para su historia y debate*. México CESU / Plaza y Valdés, UNAM (Primera reimpresión, 2000)
- Díaz Barriga, A. (1993c) "El problema de la teoría de la evaluación y cuantificación del aprendizaje" En *El examen, textos para su historia y debate*. México CESU / Plaza y Valdés/ UNAM (Primera reimpresión, 2000)
- Díaz Barriga, A. (1997) "Tesis para una teoría de la evaluación sus derivaciones en la docencia" En *Didáctica y Currículum*. México Paidós (Edición corregida y aumentada)
- Díaz Barriga, A. (2004) "Comentarios de la mesa redonda Comparabilidad de los resultados de las evaluaciones" En INEE *Jornadas de Evaluación Educativa. Memoria*. México INEE
- Díaz Barriga, A. (2005a) "El examen". En *El docente y los programas escolares. Lo institucional y lo didáctico*. Barcelona- México, Pomares.
- Díaz Barriga, A. (2005b) "Diversidad Cultural y currículum ¿Es factible una articulación?" Chile *Pensamiento Educativo*. Vol 37 pp 52-63
- Díaz Barriga, A. (2006) "Las pruebas masivas Análisis de sus diferencias técnicas" En *Revista Mexicana de Investigación Educativa*. Vol 11, No 29 pp 583-615
- Díaz Barriga, A. (2007 diciembre) "Observaciones sobre PISA" *Suplemento universitario Campus Milenio*. No 254. Recuperado diciembre 2007, de <http://www.campusmilenio.com.mx/254/ensayos/pisa.php>



- Díaz Barriga, A (2008, 22 de mayo) "Alianza por la Educación" Una remembranza de éxitos y fracasos *Suplemento Universitario Campus Milenio*. Recuperado 8 de junio de 2008, de <http://www.campusmilenio.com.mx/273/ensayos/alianza.php>
- Díaz Barriga, A et al (2009) *Reforma integral de la Educación Básica 2009. Diplomado para maestros de primaria. Módulo 1. Materiales anexos*. México: SEP, SNTE, SEB.UNAM/IIISUE.
- Díaz Barriga, F y Hernández, G (2002) *Estrategias docentes para un aprendizaje significativo* Una interpretación constructivista México Mc Graw Hill
- Díaz Barriga Arceo, F (2005) "La evaluación auténtica centrada en el desempeño una alternativa para evaluar el aprendizaje y la enseñanza" En *Enseñanza situada, vínculo entra la escuela y la vida*. México Mc Graw Hill
- Durkheim (s/f) "Los grados", tomado de *Historia de la Educación y de las doctrinas pedagógicas*. Madrid La piqueta
- Ebel, R (1977) "La medición educacional Perspectiva histórica" En *Fundamentos de la Medición Educacional*. Buenos Aires, Argentina, Editorial Guadalupe En Díaz Barriga, A. (2000) *El examen, textos para su historia y debate*. México CESUUNAM
- Escamilla, G (2005-2006) "Los resultados del concurso de ingreso a bachillerato 2005 En la ZMCM" *Revista Mexicana de Orientación Educativa*. (México), Vol 3, No 7, Mes oct-feb, Año 2005-2006, pp 12-14
- Escudero T (2003) "Desde los test hasta la investigación evaluativa actual. Un siglo, el XX, de intenso desarrollo de la evaluación en educación" En *Revista Electrónica de Investigación y Evaluación educativa* (RELIEVE), Vol , 9, No 1, Recuperado abril de 2005, de <http://www.uv.es/RELIEVE/>
- Ferrer, J G y Arregui P. (2003). *Las pruebas internacionales de aprendizaje en América Latina y su impacto en la calidad de la educación. Criterios para guiar futuras aplicaciones*. Programa de Promoción de la Reforma en América Latina y el Caribe Recuperado marzo de 2009, de [http://www.grade.org.pe/asp/brw\\_publ1.asp?id=539](http://www.grade.org.pe/asp/brw_publ1.asp?id=539)
- Ferrer, J G (2006) *Sistemas de evaluación de aprendizajes en América Latina* Balances y desafíos Santiago de Chile PREAL
- Frank, W (1997). "Evaluación educacional internacional El ejemplo de la IEA, su investigación y desafío" En C De Moura, y M Camoy, (eds) *¿Qué rumbo debe tomar el mejoramiento de la educación en América Latina?* Seminario sobre reforma educativa en América Latina Washington, D C Departamento de Desarrollo Sostenible. Banco Interamericano de Desarrollo (BID) Recuperado abril de 2005, de: <http://www.iadb.org/sds/doc/Edu%2DCCastroS.pdf>
- Frary, R T. B (1995, mayo) More multiple-choice item writing do's and don'ts *Practical Assessment, Research & Evaluation*, 4(11) Recuperado mayo de 2008, de <http://PAREonline.net/getvn.asp?v=4&n=11>

- Fontanive, N. (2005) "O uso pedagógico dos testes" En Alberto de Mello e Souza (org) *Dimensões da Avaliação Educacional*. Petrópolis, Brazil: Vozes
- Foucault, M., (1977) *Vigilar y castigar*. México: Siglo XXI
- García Cortés, F. (1979) *Paquete de autoenseñanza de evaluación del aprovechamiento escolar*. México: CISE/UNAM
- García, E. y Fidalgo, A. (2005). "Análisis de los ítems" En J. Muñiz, A. Fidalgo, E. García R. Martínez, y R. Moreno, R. *Análisis de los ítems*. Madrid España: La muralla
- Gil, M. (2004) Comentarios a la mesa "Comparabilidad de los resultados de las Evaluaciones", por Felipe Martínez Rizo. En INEE. *Jornadas de Evaluación Educativa*. Memoria México: INEE
- Gil, M. (2005, noviembre) "¿Qué hacer con los rankings?" *La Crónica Hoy*. Recuperado 28 de noviembre de 2005, de <http://irgoa.usc.es/drupal/taxonomy/term/325>
- Gil, M., (2010) Puntuales y tercas tozudas. *Revista Educación No. 9* México: Suplemento de la Universidad Autónoma de la Ciudad de México (UACM)
- Giner de los Ríos, F. (s/f) "O educación o exámenes" En A. Díaz Barriga (comp), (1993) *El examen, textos para su historia y debate*. México: CESU Plaza y Valdés/ UNAM (Primera reimpresión, 2000)
- Gómez, B. M., (1990) *Evaluación criterial. Una metodología útil para diagnosticar el nivel de aprendizaje de los alumnos*. Madrid: Narcea
- Gómez, J. y Hernández, A. (comp) (1991) *El debate social en torno a la educación. Enfoques predominantes*: Antologías
- Haladyna, T. M. (2004) *Developing and Validating Multiple-choice Items*. (3a ed.) EE.UU: Lea
- Hambleton, R. y Rodgers, J. (1995) Item bias review. *Practical Assessment, Research & Evaluation*, 4(6). Recuperado mayo de 2008, de <http://PAREonline.net/oaetvn.asp?v=4&n=6>
- Herbart, J. F. (1806-1983) *Pedagogía general derivada del fin de la educación*. España: Humanitas
- Hernández Ruiz, S. (1949) "Espíritu de las pruebas de promoción. El problema de los exámenes" En *Metodología General de la Enseñanza*. Tomo 2
- Hernández Ruiz, S. (1972) "Exámenes y exámenes" En *La crisis de la educación contemporánea*. Madrid: Compañía Bibliográfica Española



- Instituto Latinoamericano de Comunicación Educativa (ILCE) (1993) "Tecnología Educativa. Apuntes sobre su campo de acción" México. Dirección de Investigación y comunicación Educativa. Recuperado abril de 2005, de: [http://investigacion.ilce.edu.mx/panel\\_control/doc/tyc19.pdf](http://investigacion.ilce.edu.mx/panel_control/doc/tyc19.pdf)
- Instituto Nacional para la Evaluación de la Educación (2008) *PISA en el Aula: Lectura*. México, INEE
- James, W (1975) *Pragmatismo* (Tr. Lus Rodríguez Aranda) Buenos Aires, Argentina. Aguilar
- Jurado Valencia F., (coord) (2009) *Los sistemas nacionales de evaluación en América Latina: ¿Impacto pedagógico u obediencia institucional?* Universidad Nacional de Colombia. Bogotá. Facultad de Ciencias Humanas (Colección Biblioteca General. Biblioteca Abierta)
- Kamin, L (1983) *Ciencia y política del cociente intelectual*. Madrid. Siglo XXI
- Kant E (1979) *Filosofía de la historia*. (3ª reimpresión 2000). México. Fondo de Cultura Económica
- Kilpatrick, W H Harold, R, Waehburne, C y Bonser, F G (1944) *El Nuevo programa escolar*. Buenos Aires, Argentina: Losada.
- Kilpatrick, W H (1968) "Vivir y aprender una concepción nueva y vieja del proceso de aprender" En, *La función social, cultural y docente de la escuela*. Buenos Aires. Losada
- Labrador, C et al (1986) *La ratio studiorum de los jesuitas*. Traducción al castellano. Madrid. Universidad Pontificia. Comillas de Madrid.
- Leymonié, J (2008) "Nuevas formas de enseñar, nuevas formas de evaluar" En *Páginas de educación*. Año 1, No 1. Uruguay. Universidad Católica
- Laboratorio Latinoamericano de la Evaluación de la Calidad de la Educación (2008) "Reflexiones en torno a la calidad educativa en América Latina y el Caribe" Santiago, Chile. OREAL / UNESCO/ LLECE. Recuperado junio de 2008 de: <http://www.oei.es/0608.htm>
- Malo, S (s/f) "Informe de labores 2002-2005". *Boletín*. México. Centro Nacional de Evaluación para la Educación Superior (CENEVAL)
- Martínez Rizo, F (2001) "Evaluación educativa y pruebas estandarizadas. Elementos para enriquecer el debate" *Revista de la educación superior*. Vol. XXX (4), No. 120
- Martínez, R y Moreno, R (2002) "Integración de teoría sustantiva, diseño de pruebas y modelos de análisis de medición psicológica" En A. Bazán y A. J. Arce (eds) *Estrategias de evaluación y medición del comportamiento en psicología*. México. Instituto Tecnológico de Sonora. Universidad Autónoma de Yucatán
- Martínez, R, Moreno, R y Muñiz, J (2005) "Construcción de los ítems" En: J. Muñiz, A. Fidalgo, E. García Cueto, R. Martínez, R. y R. Moreno, *Análisis de los ítems*. Madrid. Pirámide

- Marzano, R. J (2001) *Designing a new taxonomy of educational objectives*. Thousand Oaks, CA: Cronwin Press
- Mauri, T y Miras, M (1996) *La evaluación en el centro escolar*. Barcelona- España: Instituto de Ciencias de la Educación de la Universidad de Barcelona: Grao
- Mauri, T (2007) *Evaluación, autorregulación y proceso de enseñanza y aprendizaje. Hacia una evaluación formativa, continua y auténtica*. Madrid: Ministerio de Educación y Ciencia.
- Mehrens, W. A (1989 Septiembre) Preparing students to take standardized achievement tests. *Practical Assessment, Research & Evaluation*, 1(11). Recuperado Mayo 3, 2008 de <http://PAREonline.net/getvn.asp?v=1&n=11>
- Meirieu, P (1992) *Aprender sí. Pero ¿Cómo?* España: Octaedro
- Meirieu, P (2004) "La clase, referencias para una práctica" En *La escuela hoy*. España: Octaedro
- Meza de Vernet, I (2002) *Procesos cognitivos Básicos*. UNIMET (Presentación en power point) recuperado mayo de 2009 de <http://ares.unimet.edu.ve/programacion/mqr01/material%20paq%20web%20vieja/Procesos%20Cognitivos%20Basicos.ppt#7>
- Moreno, G, Huerta, C y Castellanos, J. D (coord) (2007) *Aprobado. Guía de estudio para tus evaluaciones*. México: Santillana
- Moskal, B. M. y Leydens, J. A (2003) "Scoring rubric development: validity and reliability" *Practical Assessment, Research & Evaluation*, pp 7-10. Recuperado en mayo de 2008, de <http://PAREonline.net/getvn.asp?v=7&n=10>
- Mouchpn, S (2005) "Contribuições da psicologia cognitiva à avaliação". En: Alberto de Mello e Souza (org) *Dimensões da Avaliação educacional*. Petrópolis, Brazil: Vozes
- Muñiz, J (1997) *Introducción a la Teoría de respuesta a los Items*. Madrid: Pirámide.
- Muñiz, J, Fidalgo, A, García E, Martínez, R y Moreno, R (2005) *Análisis de los Items*. Madrid: España: La muralla
- Muñoz Izquierdo, C (2008, junio) "Alianza por la Calidad de la Educación" *Observatorio ciudadano de la educación*. Artículos de opinión. Recuperado noviembre de 2008, de [http://www.observatorio.org/opinion/AlianzaCalidad\\_CMI.html](http://www.observatorio.org/opinion/AlianzaCalidad_CMI.html)
- Nichols, S y Berliner, D. (2007). *Collateral Damage. How High-Stakes Testing Corrupts America's Schools*. Cambridge: Harvard Education Press.
- Organisation, for Economic Co-operation and Development, *PISA 2003 Technical Report*, Recuperado 2005, de <http://www.pisa.oecd.org/dataoecd/4/9/6035188570.pdf>

- OCDE, (2000) *Proyecto PISA: La medida de los conocimientos y destrezas de los alumnos: un nuevo marco de evaluación*. OCDE — Madrid Ministerio de Educación, Cultura y Deporte, INCE Recuperado junio de 2008, de <http://www.ince.mec.es/pub/pisacomp.pdf> (Original publicado en 1999)
- OCDE, (2001) *Proyecto PISA La medida de los conocimientos y destrezas de los alumnos: la evaluación de la lectura, las matemáticas y las ciencias en el proyecto Pisa 2000* / OCDE — Madrid Ministerio de Educación, Cultura y Deporte, INCE Recuperado agosto de 2008, de <http://www.ince.mec.es/pub/pisa2000assessment.pdf> (Original publicado en 2000)
- OCDE, (2002) *Conocimientos y aptitudes para la vida. Resultados de PISA 2000*. (Traducido por Claudia Estévez) México OCDE- Santillana Aula XXI (Original publicado en 2001)
- OCDE, (2006) *PISA 2006. Marco de la evaluación. Conocimientos y habilidades en ciencias, matemáticas y lectura*. OCDE Recuperado septiembre de 2007, de <http://www.oecd.org/dataoecd/59/2/39732471.pdf>
- Padilla, R. A. (2005) "Exámenes masivos Un análisis entre pruebas nacionales e internacionales". *Tesis de maestría* no publicada, México Universidad Nacional Autónoma de México. Facultad de Filosofía y Letras
- Pavón, M. F. (2008). "Carrera Magisterial dejará de reconocer factores que no estén alineados con la calidad y el desempeño" Entrevista con el Dr. Jorge Santibáñez Director de la Unidad de Planeación y Evaluación de Políticas educativas " *Educación 2001*, México.
- Perrenoud, P. (2002) *Construir competencias desde la escuela* Santiago de Chile Océano
- Perrenoud, P. (2008) *La evaluación de los alumnos. De la producción de la excelencia a la regulación de los aprendizajes. Entre dos lógicas* Buenos Aires, Argentina Alternativa pedagógica
- Pérez Gómez, A. I. (2000) "El aprendizaje escolar de la didáctica operatoria a la reconstrucción de la cultura en el aula" En J. Gimeno Sacristán y A. I. Pérez Gómez *Comprender y transformar la enseñanza*. (11° Ed.) Morata España.
- Picaroni, B. (2009) "Las evaluaciones en las aulas de primaria Usos formativos, calificaciones y comunicación con los padres" Programa de la Reforma Educativa en América Latina (PREAL) Recuperado 15 abril del 2010, de [http://www.preal.org/Biblioteca.asp?Id\\_Carpeta=225&Camino=315|Grupos%20de%20Trabajo%20Evaluacion%20y%20Estimacion%20de%20Estrategias%20de%20Aprendizaje%20y%20Evaluacion%20de%20Aprendizaje](http://www.preal.org/Biblioteca.asp?Id_Carpeta=225&Camino=315|Grupos%20de%20Trabajo%20Evaluacion%20y%20Estimacion%20de%20Estrategias%20de%20Aprendizaje%20y%20Evaluacion%20de%20Aprendizaje)
- Pozo, I. (2002) *Aprendices y Maestros. La nueva cultura del aprendizaje*. Madrid Alianza Editorial
- Poy, L. (2008, 9 de abril) "A partir del día 14 se aplicará la prueba Enlace en escuelas primarias SEP" *La Jornada*, Recuperado el 24 de octubre de 2008, de <http://www.jornada.unam.mx/2008/04/09/index.php?section=sociedad&article=042n1soc>.

- Rama, G. (1994) "Necesidades de información para la toma de decisiones en una gestión moderna de los ministerios de educación" *Boletín del Proyecto Principal de Educación en América Latina y el Caribe*. (Chile) No 33, abril del 1994, pp 41-46
- Ramírez, R., (2010) "La prueba Enlace Contra el sentido de la Educación" *Revista Educación No. 9*. México: Suplemento de la Universidad Autónoma de la Ciudad de México (UACM)
- Ravela, P. (2006a) *Para comprender las evaluaciones educativas. Fichas didácticas*. Grupo de trabajo sobre Evaluación y Estándares e América Latina y el Caribe (GRADE). Programa de promoción de la Reforma Educativa en América Latina y el Caribe (PREAL) (Biblioteca 1 Estudios encargados) Recuperado 19 de noviembre de 2007, de [http://www.preal.org/Biblioteca.asp?Id\\_Carpeta=225&Camino=315%7CGRUPOS%20DE%20TRABAJOS%20Y%20ESTANDARES%20Y%20PUBLICACIONES](http://www.preal.org/Biblioteca.asp?Id_Carpeta=225&Camino=315%7CGRUPOS%20DE%20TRABAJOS%20Y%20ESTANDARES%20Y%20PUBLICACIONES).
- Ravela, P. (2006b) "¿Qué son las evaluaciones y para qué sirve? Un recorrido por las evaluaciones y sus finalidades" Ficha No 1 En *Para comprender las evaluaciones educativas. Fichas didácticas*. Grupo de trabajo sobre Evaluación y Estándares e América Latina y el Caribe (GRADE) Programa de promoción de la Reforma Educativa en América Latina y el Caribe (PREAL) (Biblioteca 1 Estudios encargados) Recuperado 19 de noviembre de 2007, de <http://www.preal.org/GTEE/doc/FichDid%20-%20Ficha%2001.pdf>
- Ravela, P. (2006c) "¿Cómo se hacen las evaluaciones educativas? Los elementos básicos del proceso de evaluación" Ficha No 2 En *Para comprender las evaluaciones educativas. Fichas didácticas*. Grupo de trabajo sobre Evaluación y estándares e América Latina y el Caribe (GRADE) Programa de promoción de la Reforma Educativa en América Latina y el Caribe (PREAL) (Biblioteca 1 Estudios encargados) Recuperado 19 de noviembre de 2007, de <http://www.preal.org/GTEE/doc/FichDid%20-%20Ficha%2002.pdf>
- Ravela, P. (2006d) "¿Cómo se formulan los juicios de valor? ¿Evaluaciones normativas, de progreso y criterios?" Ficha No 3 En *Para comprender las evaluaciones educativas. Fichas didácticas*. Grupo de trabajo sobre Evaluación y estándares e América Latina y el Caribe (GRADE) Programa de promoción de la Reforma Educativa en América Latina y el Caribe (PREAL) Recuperado el 7 de noviembre de 2007, de <http://www.preal.org/GTEE/doc/FichDid%20-%20Ficha%2003.pdf>
- Ravela, P. (2006e). "¿Cuáles son los principales problemas comunes a todas las evaluaciones? Validez y confiabilidad" Ficha No 4 En *Para comprender las evaluaciones educativas. Fichas didácticas*. Grupo de trabajo sobre Evaluación y estándares e América Latina y el Caribe (GRADE). Programa de promoción de la Reforma Educativa en América Latina y el Caribe (PREAL) (Biblioteca 1 Estudios encargados) Recuperado 19 de noviembre de 2007, de <http://www.preal.org/GTEE/doc/FichDid%20-%20Ficha%2004.pdf>
- Ravela, P. (2006f) "¿Qué evalúa esta prueba? I. Distintos tipos de actividades en las evaluaciones estandarizadas" Ficha No 6 En *Para comprender las evaluaciones educativas. Fichas didácticas*. Grupo de trabajo sobre Evaluación y estándares e América Latina y el Caribe (GRADE) Programa de promoción de la Reforma Educativa en América Latina y el Caribe (PREAL) (Biblioteca 1 Estudios encargados) Recuperado el 19 de noviembre de 2007, de <http://www.preal.org/GTEE/doc/FichDid%20-%20Ficha%2005.pdf>.

- Ravela, P. (2006g) ¿Qué evalúa esta prueba? II. Contenidos, currículo y competencias. Ficha No. 7. En *Para comprender las evaluaciones educativas. Fichas didácticas*. Grupo de trabajo sobre Evaluación y estándares e América Latina y el Caribe (GRADE). Programa de promoción de la Reforma Educativa en América Latina y el Caribe (PREAL) (Biblioteca 1 Estudios encargados). Recuperado el 19 de noviembre de 2007, de <http://www.preal.org/GTEE/docr/FichDid%20-%20Ficha%2005.pdf>.
- Ravela, P. (2006h) ¿Qué significan los números de las evaluaciones? I. Elementos básicos para comprender los datos estadísticos. Ficha No. 8. En *Para comprender las evaluaciones educativas. Fichas didácticas*. Grupo de trabajo sobre Evaluación y estándares e América Latina y el Caribe (GRADE). Programa de promoción de la Reforma Educativa en América Latina y el Caribe (PREAL) (Biblioteca 1 Estudios encargados). Recuperado el 19 de noviembre de 2007, de <http://www.preal.org/GTEE/docr/FichDid%20-%20Ficha%2005.pdf>.
- Ravela, P. (2006i) ¿Por qué los rankings son modos inapropiados de valorar la calidad de las escuelas? Hacia nuevos modos de utilizar los resultados por escuelas. Ficha No. 10. En *Para comprender las evaluaciones educativas. Fichas didácticas*. Grupo de trabajo sobre Evaluación y estándares e América Latina y el Caribe (GRADE). Programa de promoción de la Reforma Educativa en América Latina y el Caribe (PREAL) (Biblioteca 1. Estudios encargados). Recuperado 19 de noviembre de 2007, de <http://www.preal.org/GTEE/docr/FichDid%20-%20Ficha%2005.pdf>.
- Ravela, P. (2006j) ¿Qué es la rendición de cuentas? Hacia la responsabilidad compartida. Ficha No. 11. En *Para comprender las evaluaciones educativas. Fichas didácticas*. Grupo de trabajo sobre Evaluación y estándares e América Latina y el Caribe (GRADE). Programa de promoción de la Reforma Educativa en América Latina y el Caribe (PREAL) (Biblioteca 1 Estudios encargados). Recuperado 19 de noviembre de 2007. <http://www.preal.org/GTEE/docr/FichDid%20-%20Ficha%2005.pdf>.
- Ravela, P. (2006k) ¿Qué son los factores asociados? Intentando comprender los sistemas educativos. Ficha No. 13. En *Para comprender las evaluaciones educativas. Fichas didácticas*. Grupo de trabajo sobre Evaluación y estándares e América Latina y el Caribe (GRADE). Programa de promoción de la Reforma Educativa en América Latina y el Caribe (PREAL) (Biblioteca 1. Estudios encargados). Recuperado el noviembre de 2007, de <http://www.preal.org/GTEE/docr/FichDid%20-%20Ficha%2005.pdf>.
- Ravela, P. (2006l) ¿Cómo analizar un reporte de evaluación? Las preguntas que el lector debe hacerse ante un informe de resultados. Ficha No. 14. En *Para comprender las evaluaciones educativas. Fichas didácticas*. Grupo de trabajo sobre Evaluación y estándares e América Latina y el Caribe (GRADE). Programa de promoción de la Reforma Educativa en América Latina y el Caribe (PREAL) (Biblioteca 1 Estudios encargados). Recuperado el 19 de noviembre de 2007, de <http://www.preal.org/GTEE/docr/FichDid%20-%20Ficha%2005.pdf>.

- Ravela, P (2007) "La evaluación como motor de cambio en la educación" *Lección inaugural*. Universidad Católica de Uruguay. Recuperado diciembre de 2009, de [http://www.uca.edu.uy/Portals/0/Publico/Facultades/Ciencias%20Humanas/IEE/leccion inaugural 2007-pravela.pdf](http://www.uca.edu.uy/Portals/0/Publico/Facultades/Ciencias%20Humanas/IEE/leccion%20inaugural%2007-pravela.pdf).
- Ravela, P, Arregui, P, Valverde G, Wolfe, R, Ferrer, G, Martínez Rizo, F, Aylwin, M Y Wolff, L (2008) "Las evaluaciones educativas que América Latina necesita" No 40 Santiago Chile PREAL. Recuperado 9 de marzo del 2010 de: [http://www.preal.org/Biblioteca.asp?Id\\_Carpeta=64&Camino=63%7CPreal%20Publicaciones%64%7CPREAL%20Documentos](http://www.preal.org/Biblioteca.asp?Id_Carpeta=64&Camino=63%7CPreal%20Publicaciones%64%7CPREAL%20Documentos)
- Romano, L (2003, 13 de agosto) "Dan palomita a secundarias capitalinas" *Reforma*. México
- Romano, L (2003, 13 de agosto) "Publica CENEVAL cuadro de honor" *Reforma*. México
- Salvat H (1978) "La escuela una carrera de obstáculos" En L Sève, M Verret, y G Snyders *El fracaso escolar*. México Ediciones de cultura popular.
- Sánchez, M (2001) *Cuarto análisis de la prueba censal en matemática. Estudio de detección de errores persistentes en Matemática*. Montevideo, Uruguay ANEP-CODIGEN-DIE
- Sánchez, M (2008a) *La evaluación PISA y su impacto en Uruguay*. Congreso Internacional de Evaluación Educativa Tlaxcala, México (Presentación en power point)
- Sánchez, M (2008b) *Programa PISA en Uruguay*. Jornada de trabajo con docentes (Presentación en power point)
- Sanders W y Horn, S (1995), "Educational Assessment Reassessed The Usefulness of Standardized and Alternative Measures of Student Achievement as Indicators for the Assessment of Educational Outcomes" *Education Policy Analysis Archives* Vol\_ 3 No\_ 6 Sanders & Horn Educational Assessment Reassessed
- Schiefelbein, E et al (1995). "Calidad de la educación, desarrollo, equidad y pobreza en la región, 1980-1994" *Boletín: Proyecto principal de educación en América Latina y el Caribe*. (CHILE), No 38, Mes Dic., Año 95, pp 3-51.
- Snyders G (1972) *Pedagogía progresista*. Educación tradicional y educación nueva. Madrid. Marova
- SEP (1993) *Plan y Programas de Estudios 1993. Educación Básica Primaria*. México: SEP
- SEP (2002) *La experiencia de la Dirección General de Evaluación en la Educación Básica y Normal*. Recuperado septiembre de 2008, de: <http://www.sep.gob.mx/work/apps/site/dge/index.htm>
- SEP (2007) "ENLACE Educación básica. Elaboración de reactivos" Recuperado 18 de noviembre de 2008, de: <http://enlace2007.sep.gob.mx/content/view/full/112/>



- SEP (2007) *Programa Sectoral de Educación 2007-2012*. Documento recuperado en el portal SEP 2 de enero de 2008, de: [http://upepe.sep.gob.mx/prog\\_sec.pdf](http://upepe.sep.gob.mx/prog_sec.pdf)
- SEP (2008a) "Alianza Por la calidad de la educación" Boletín No. 119. Recuperado 16 de mayo 2008, de <http://www.sep.gob.mx/wb/sep1/bol1190508>.
- SEP (2008b) *Plan de estudios 2009. Educación Básica Primaria*. México: SEP
- SEP (2008c) Evaluación Nacional del Logro Académico en Centros Escolares. *Sugerencias para el uso pedagógico de los resultados de ENLACE. Sexto de primaria*.
- SEP (s/f) *Análisis de ítems. Índice de dificultad. Índice de discriminación*. Versión en power point
- SEP (s/f) Evaluación Nacional del Logro Académico en Centros Escolares, ENLACE Educación básica. Distribución de reactivos de Español por tipo de texto y línea de evaluación. Sexto Grado de Primaria. Recuperado 7 de noviembre del 2009 de [http://enlace.sep.gob.mx/ba/viewimg.php?img=img/tablas/esp\\_p6.png](http://enlace.sep.gob.mx/ba/viewimg.php?img=img/tablas/esp_p6.png).
- Solano, L. (2005, junio). "Hugo Aboites afirma: Examen único excluye a los jóvenes de la educación" *La jornada*. Recuperado septiembre de 2009, de: <http://www.educnm.org/ruqa/612/site/default/files/Aboites.pdf>.
- Solano, G. y Backhoff, E. (2003). *La traducción de pruebas en comparaciones internacionales: México*. INEE. Recuperado abril de 2008, de [http://www.inee.edu.mx/images/stories/documentos\\_pdf/Resultados\\_Evaluaciones/Estudios/Anteriores\\_2003/07\\_aspects\\_lingu\\_traductimss95.pdf](http://www.inee.edu.mx/images/stories/documentos_pdf/Resultados_Evaluaciones/Estudios/Anteriores_2003/07_aspects_lingu_traductimss95.pdf).
- Solano- Flores, G., Contreras -Niño, L. A. y Backhoff, E. (2006) "Traducción y adaptación de pruebas: Lecciones aprendidas y recomendaciones para países participantes en TIMSS, PISA y otras comparaciones internacionales" En *Revista Electrónica de Investigación Educativa, REDIE*. Vol.8, No. 2. México: Ensenada, Baja California. Recuperado junio de 2008, de: <http://redie.uabc.mx/>
- Stufflebeam, D. y Shinkfield, A. J. (1987). *Evaluación Sistemática. Guía teórica y práctica* (Tr. Carlos Losilla). Barcelona- México: Paidós.
- Székely, B. (1966) *Los test*. (5ª Ed.) Argentina: Kapeluz.
- Taba, H. (1979) *Elaboración del currículo* (4ª Ed.) Argentina: Troquel.
- Terman, L. M. (1916) *The measurement of intelligence: an explanation of and a complete guide for the use of the Stanford revision and extension of the Binet-Simon intelligence scale*. Boston: Houghton Mifflin Company.
- Thorndike, R. y Hagen, E. (1978) "Orientación histórica y filosófica" En *Tests y técnicas de medición en psicología y educación*. México: Trillas.

- Tiana, A (1996) "La evaluación de los sistemas educativos". *Revista Iberoamericana de Educación. No. 10*. Madrid Biblioteca de la Organización de Estados Iberoamericanos para la Educación, la Ciencia y la Cultura Recuperado en mayo de 2004, de <http://dialnet.unirioja.es/servlet/oaiart?codigo=1019993>
- Tiana, A (1997) *Tratamiento y uso de la información en evaluación* Organización de Estados Iberoamericanos para la Educación, la Ciencia y la Cultura (OEI) Recuperado junio de 2005, de <http://www.oei.org.ar/noticias/tratamiento.pdf>.
- Tristán, A (1998) *Análisis de Rasch para todos. Guía simplificada para evaluadores educativos*. México CENEVAL
- Tort, M (1979) "La situación de un test Una situación social" En *El cociente intelectual* México Siglo XXI
- Tyler, R (1967) "Examen y valoración del conocimiento, destreza y capacidad adquiridos" En F Freeman, *La pedagogía científica* (1967) Buenos Aires, Argentina Losada
- UNESCO (2001) *Balance de los 20 años del Proyecto Principal de Educación en América Latina y el Caribe. Chile: Oficina Regional de Educación de la UNESCO para América Latina y el Caribe OREALC/UNESCO* Recuperado mayo de 2005, de <http://www.unesco.cl/promedi7/balance.pdf>
- Valdés, H (coord) (2008) *Segundo Estudio Regional Comparativo y Explicativo (SERCE). Los aprendizajes de los estudiantes de América Latina y el Caribe* Resumen ejecutivo Santiago de Chile. LLECE / UNESCO Recuperado 12 de marzo de 2009, de <http://unesdoc.unesco.org/images/0016/001606/160659s.pdf>
- Velloso, A. y Pedró, F. (1991). *Manual de educación comparada*. Vol 1 Barcelona, España Universitat
- Vidal, R. (2008). La evaluación Educativa. ¿ENLACE; EXANI, EXCALE o PISA? (primera de tres partes de tres partes) Recuperado 2 de julio de 2009, de <http://www.campusmilenio.com.mx/327/ensayos/evaluacion.php>.
- Wildemuth, B (1984) *Alternatives to Standardized Tests*. Recuperado el 11 de octubre de 2005, de: [http://www.eric.ed.gov/ERICDocs/data/ericdocs2/content\\_storage\\_01/00000000/b80/2a/0b/da.pdf](http://www.eric.ed.gov/ERICDocs/data/ericdocs2/content_storage_01/00000000/b80/2a/0b/da.pdf)
- Witrock, M y Baker, E. (1998) *Test y cognición. Investigación cognitiva y mejora de las pruebas psicológicas*. España Paidós
- Wood, A., (1971) *Elaboración de tests. Desarrollo e interpretación de los test de aprovechamiento* México, Trillas (Título de la obra en inglés test constructions, 1960)
- Wolff, L (2007) "Los costos de las evaluaciones de aprendizaje en América Latina" No 38 Washington, D.C. PREAL / GRADE. Recuperado mayo de 2010, de [http://www.prea.org/Biblioteca.asp?Id\\_Carpeta=225&Camino=315%7CGrupos%20de%20Trabajo%88%7CEvaluaci%F3n%20v%20Es:%E1ndares/225%7CPublicaciones](http://www.prea.org/Biblioteca.asp?Id_Carpeta=225&Camino=315%7CGrupos%20de%20Trabajo%88%7CEvaluaci%F3n%20v%20Es:%E1ndares/225%7CPublicaciones)